

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

## Om nucleic - protein search, using frame\_plus\_n2p model

Run on: January 31, 2003, 06:13:31 ; Search time 49.5 Seconds

(without alignments)  
6418.680 Million cell updates/sec

Title: US-09-598-982-20  
Perfect score: 1458  
Sequence: 1 gggccctcgagaaaaaat.....cgtgaagcgccgcgtcg 771

## Scoring table:

BLOSUM62

Xrapop 10.0	Xapext 0.5
Xgpop 10.0	Xapext 0.5
Frapop 6.0	Fgapext 7.0
Delop 6.0	Dejext 7.0

Searched: 671580 seqs, 205047115 residues

Total number of hits satisfying chosen parameters: 1343160  
Minimum DB seq length: 0  
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-KODEL-frame-n2p.model -DEV=xlh

```
-O=/cgn1/1/USPTO_spool/rus9892/runat_27012003_073716-19567/app-query.fasta_1.967
-DB=SPTREMBL_21_QFM=fastan -SUFFIX=n2p_rprt -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-O=CGN1/1/USPTO_spool/rus9892/runat_27012003_073716-19567/app-query.fasta_1.967
-MODE=LOCAL -OUTfmt PRO_NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09598-982-0@runat_27012003_073716-19567 -NCPU=6 -TCPU=3
-NODUMP -NO_MMAP -LARGE_MODE -NEUT_SCORING=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database : SPTREMBL\_21:\*

```
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_orcanelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_primate:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacterioplasm:*
17: sp_archeap:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match Length	DB ID	Description
1	1342	92.0	275	4 Q96RZ6 homo sapien

## SUMMARIES

## ALIGNMENTS

```
RESULT 1
Q96RZ6 PRELIMINARY; PRT; 275 AA.
ID Q96RZ6
AC Q96RZ6;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
TRYPSE T.
GN TRYPSEC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9605;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=21096910; PubMed=11157797;
RX Daniels R.J., Peden J.F., Lloyd C., Horsley S.W., Clark K., Flint J., Tufarelli C., Kearney L., Buckle V.J., Doggett N.A., Flint J., Higgs D.R.; "Sequence, structure and pathology of the fully annotated terminal 2 Mb of the short arm of human chromosome 16."; Hum. Mol. Genet. 10:339-352(2001).
RL DR EMBL: AE06465; ARK61271.1; MEROPS: S01.242; InterPro: IP001254; Ser_protease_Try.
```

Pfam: PF00089; trypsin; 1.  
 DR PROSITE; PS50240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
 DR PROSITE; PS00155; TRYPSIN\_SER; UNKNOWN\_1.  
 KW Hydrolase; Serine protease.  
 SEQUENCE 275 AA; 30601 MW; 2F4AAAETDDE483E CRC64;  
 Alignment Scores:  
 Pred. No.: 1.54e-94 Length: 275  
 Score: 1342.00 Matches: 240  
 Percent Similarity: 97.96% Conservative: 0  
 Best Local Similarity: 97.96% Mismatches: 5  
 Query Match: 92.04% Indels: 0  
 DB: 4 Gaps: 0  
 US-09-598-982-20 (1-771) x Q96R26 (1-275)  
 QY 19 ATCGTGGGGTCAGGAGCCCCAGGAGRAAGTGGCCCTGCAGGTGAGAGTC 78  
 Db - 31 IleValGlyGlyGlnGluAlaProArgSerIysTrpProTrpGlnValSerIeuArgVal 50  
 QY 79 CACGGCCATACTGGATGCACTCTCGGGAGCTCCTCACCCACGGGAGCTG 138  
 Db 51 HISAsProTyTrpMetHisPheCysGlyGlySerIeuIleHisProGlnTrpValLeu 70  
 QY 139 ACCGCCGCCGCGTGGCGGGAGCTGAGCTGGCCCTCAAGTGACT 198  
 Db 71 ThrAlaLahisCysValGlyProAspIysValLeuAspIeuAlaLeuArgValGlnLeu 90  
 QY 199 CGGGAGGAGCACCTCTAACCGAGCACGCTGCTGCCGTCAGCAGCATCGGCCAC 258  
 Db 91 ArgGluGlnHisLeuTyTrpGlnAspGlnLeuLeuProValSerArgIleIleValHis 110  
 QY 259 CCACAGTCTACACGCCAGATCGGGAGCATCCCTCTGGAGCTGGAGGCC 318  
 Db 111 ProGlnPheTyTrhAlaGlnIleGlyAlaAspIleAlaLeuGluGluIlePro 130  
 QY 319 GNGAAGSTCTCAGCCAGCTCACAGCTTCCGCTTCCAGAGACTTCCC 378  
 Db 131 ValAsnValSerSerHisValHistnValThrLeuProProLaserGluThrPhePro 150  
 QY 379 CGGGGTGCGCTGCGGGTCACTGGCAATGATGAGCAGCCCTCCA 438  
 Db 151 ProGlyMetProCystTrpValThrGlyLysTrpGlyAspIysValAspAsnAspGluArgLeupro 170  
 QY 439 CGCCCATTTCTCTGAAGCAGGTGGAACTCCATAATGAAACCATTTGTGACCA 498  
 Db 171 ProProHeProLeuIysGlnValIysValProIleMetGluAsnHisIleCysAspIala 190  
 QY 499 AAATACCACTTGGCCTACAGGGAGGAGCTCCGCTGACACATCTG 558  
 Db 191 LysTyRHisLeuGlyIalatYThrGlyAspAspValLargIleValArgAspAspMetIeu 210  
 QY 559 TGTGCGCCGGAAACCCGAGGAACTCTGGAGGCCCTGGTGC 618  
 Db 211 CysAlaGlyAsnThrArgArgAspSerGlyGlyProLeuValCys 230  
 QY 619 AAGGTGAATGGACCTGGCTGCGGGGGTGTCACTGGGGAGGGCTGGCCAG 678  
 Db 231 LysValAsnGlyThrTrpIleGlnAlaGlyValArgTrpGlyGlyIysAlaGln 250  
 QY 679 CCGAACCGGCGCTGCGCATCTACACCGGTGACTACTACTTGAGCTGACATTCCACCATAT 738  
 Db 251 ProAsnAlaGlyProGlyIleTyTrhArgValThrIrrYleuAsnAspIleHis\*\*\*Tyr 270  
 QY 739 GTCCCCAAAGAAGCG 753  
 Db 271. ValProLySlysPro 275  
 RESULT 2  
 Q921N4 PRELIMINARY; PRT; 273 AA.  
 ID Q921N4; OC OC  
 AC OC  
 DR DT 01-DEC-2001 (T:EMBLrel. 19, last sequence update)  
 DR DT 01-JUN-2002 (T:EMBLrel. 21, last annotation update)  
 DR DE Mast cell protease 7.  
 DR MCPT7.  
 KW OS muis musculus (Mouse);  
 OC OC Bokaryota; Metzoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC OC mammalia; Buteraria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;  
 OC OC NCBI\_TaxID=10090;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RA Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 RL DR EMBL; BC011328; AAH11328.1; -.  
 DR DR METOPS; S01\_026; -.  
 DR DR MGI; 96943; MCPT7.  
 DR DR InterPro; IPR001254; Ser-Protease\_Try.  
 DR DR Pfam; PF00089; trypsin; 1.  
 DR DR PROSITE; PS50240; TRYPSIN\_DOM; 1.  
 DR DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
 DR DR PROSITE; PS00155; TRYPSIN\_SER; UNKNOWN\_1.  
 KW Hydrolase; Protease; Serine protease.  
 SEQUENCE 273 AA; 30332 MW; 51417476514035BE CRC64;  
 Alignment Scores:  
 Pred. No.: 3.82e-72 Length: 273  
 Score: 1049.00 Matches: 188  
 Percent Similarity: 81.67% Conservative: 17  
 Best Local Similarity: 74.90% Mismatches: 44  
 Query Match: 71.95% Indels: 2  
 DB: 11 Gaps: 1  
 US-09-598-982-20 (1-771) x Q921N4 (1-273)  
 QY 1 GGGCCCCCTCGAGAAAGA-----ATGTCGGGGGAGGAGCCCCAGGAGCAGTGG 54  
 Db 21 GlyProAlaMetThrArgGluGlyIleValGlyGlnGluAlaIleGlyAsnIysTrp 40  
 QY 55 CCCTGGAGGGAGCCCTGAGACTTCCAGGCCATACTGGATGCACTCTGGGGGCTCC 114  
 Db 41 ProTrpGlnValSerIeuArgGalaAsnAspThrIrrPheCysGlyGlySer 60  
 QY 115 CTCATCCACCCAGGGGGTGTGAGGCCGGCGTGGGGAGGACACCTACTACGGGACAGCCTG 174  
 Db 61 LeuIleHisProGlnTrpValLeuThrAlaLahisCysValGlyProAspValAlaAsp 80  
 QY 175 CTTGGCGCCCTAGGGCAACTGCGGAGGACGACCTACTACGGGACAGCCTG 234  
 Db 81 ProAsnIysValArgValGlnIleIeuArgIysGlnTyTruIeuTyTrhHisAspHisIeuMet 100  
 QY 235 CCGCTGAGGAGTCATGTCGTCAGCCACAGTCTACACGCCAGATCGGGGACATC 294  
 Db 101 ThrValSerGinIleIleThrHisProAspPheTyTrlValGlnAspGlyAlaAspIle 120  
 QY 295 GccCTGGCGAGCTGGAGGAGGGGGAGGGCTACGGCTCACCGTCACCGTCACCGCT 354  
 Db 121 AlaIleLeuIleLeuThrAsnProValAsnIleSerAspTyValHisProValProLeu 140  
 QY 355 CCCCTGCTCAGAGACCTCCCGGGATGCCCTGGCTGCTGGCTACGGCTGGCGAT 414  
 Db 141 ProProIleAspIleThrPheProSerGlyIrrLeuIysTrpValThrGlyTrpGlyAsn 160  
 QY 415 GTGGAACATGATGAGGCCCTCCACGCCATTCCTGTAACCGAGGTGGCTCCCATAT 474  
 Db 161 IleAspGlyIysValAsnIleIeuProProProHeProLeuIysGluValGlnValProle 180  
 QY 475 ATGGAACACACATTGAGCCAAATAACACACTTGCGCTACAGGGAGCAGGTC 534  
 Db 181 IleGluAsnHisIleIeuCysAspIeuIleIrrIysIleIeuIleThrGlyAspVal 200  
 QY 535 CGGATCGCCGTCAGACATGCTGTCGCCGGAACACCCGGAGGACTCATCCGGCC 594  
 Db 201 HisIleIeuIleAspIleIeuCysAlaGlyAsnGlyIysAspSerCysGly 220



Db 29 IleIleGlyLysGluAlaProGlySerArgTrpProTrpGlnValSerLeuArgVal 48  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.  
 QY CACGGCCCATACTGGATGCACTCTGGGGGCTCCCCTATCACCCCCAGGGCTCG 138  
 DR EMAU; X94982; CA04438 1; -.  
 DR HSSP; P20231; IAAQ.  
 DR MEROPS; S01.118; -.  
 DR InterPro; IPR001314; Chymotrypsin.  
 DR Prosite; PS00240; SERPINE.  
 DR Pfam; PF00089; trypsin\_1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR SMART; SM0020; TRY\_SPEC\_1.  
 DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
 Db 49 ArgAspGlnTyrTrpArgHisGlnCysGlyLysSerLeuLeuHistProGlnTrpValLeu 68  
 DR KW Hydrolase; Serine protease.  
 DR NON\_TER 1  
 Db 139 ACCGGCCGGCGCGTGGGACCGGGACGGTCAAAGATCTGGCCACCCCTAGGGTGCAACTG 198  
 DR 69 ThrAlaAlaHisCysIleGlyLysProGluLeuGlnGluProSerAspPheArgValGlnLeu 88  
 DR InterPro; IPR001254; Ser\_protease\_Try.  
 DR Prosite; PS001314; Chymotrypsin.  
 DR PRINTS; PR00089; trypsin\_1.  
 DR SMART; SM0020; TRY\_SPEC\_1.  
 DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
 Db 199 CGGGAGCAGCACCTCTACTAACCGAGGACAGCTGTGGGGTCAGCAGGGATCATCTG 258  
 DR 89 ArgGluGlnHisLeuTyrTrpGlnAspGluLeuProIleSerArgValLeuPro 108  
 DR 259 CCACAGTCTACACGCCAGATGGAGGACATGGCTGGAAGCTGGAGGAGCCG 318  
 DR 109 ProHsTyrTyrMetTvalGluAsnGlyIaAspIleIaleLeuGlnLeuGluPro 128  
 DR 149 ProGlySerGlnCysTrpValThrGlyTrpGlyIasnValAspAsnGlyArgProLeuPro 168  
 DR 319 GTGAGGCTCTCAGCCACGGTCCACAGGTCACCTGCCCCCTGCCTCAGACGCTCCC 378  
 DR 169 ProProTyrProLeuGlyGlnValValProIleValGluAsnSerValIacySprrp 188  
 DR 499 AAATACCACCTGGCCCTACACGGAGRCGACGCTCCGCACTGGCGACGATGGC 558  
 DR 189 LysTyrIleSerGlyIleSerThrTyrSerValProIleValGlnGluAsnLeu 208  
 DR 559 TGTGCGGGGACACCCGGAGGGACTCATGCAGGGACTCTGGGGGACTGGGTG 618  
 DR 209 CysAlaGlyAspGlyGlyArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 228  
 DR 619 AAGGGAAATGCACTGGCAGCAGCGTGGCAGCTGGCTGGGGGAGGGCTGCCAG 678  
 DR 229 LysValAsnGlyIleSerThrTyrSerValProIleValGlnGlyAspGlyCysAlaAsn 248  
 DR 739 GTCCCCAAAAGCCG 753  
 DR 269 ValProGlnGluPro 273  
 RESULT 5  
 PRELIMINARY: PRT; 237 AA.  
 ID 029464 DT 01-NOV-1996 (Trembllel. 01, created)  
 AC DT 01-NOV-1996 (Trembllel. 01, last sequence update)  
 RN DT 01-JUN-2002 (Trembllel. 21, Last annotation update)  
 DE Trypsin (EC 3.4.21.59) (Fragment).  
 OS Bos taurus (Bovine).  
 OC Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Bovidae; Bovinae; Bos.  
 OC NCBI Taxon=9913;  
 RP [1] SEQUENCE FROM N A.  
 TISSUE-LIVER CAPSULE:  
 RX MEDLINE=96203914; PubMed=6620801;  
 RA Pallaro M., Gambacurta A., Florucci L., Mignogna G., Barra D.,  
 RA Ascoli F.;  
 RT "cDNA cloning and primary structure of trypsinase from bovine mast cells  
 and evidence of the expression of bovine pancreatic trypsin inhibitor  
 mRNA in the same cells.";  
 RL Eur. J. Biochem. 237:100-105(1996).

RESULT 6  
 Q96R25 PRELIMINARY; PRT; 242 AA.  
 AC Q96R25;  
 DT 01-MAR-2001 (TREMBrel. 19, last sequence update)  
 DT 01-MAR-2002 (TREMBrel. 20, last annotation update)  
 DE Putative mast cell MNCP-7-like II tryptase (MNCP-7-like-2).  
 GN MNCP-7L.  
 OS Homo sapiens (Human).  
 OC Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OX NCBI\_TAXID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21096910; PubMed=11157797;  
 RA Daniels R.J., Peden J.F., Lloyd C., Horsley S.W., Clark K.,  
 RA Tufarelli C., Kearney L., Buckle V.J., Doggett N.A., Flint J.,  
 RT Higgs D.R.;  
 "Sequence, structure and pathology of the fully annotated terminal 2 Mb of the short arm of human chromosome 16.";  
 RL Hum. Mol. Genet. 10:339-352(2001).  
 DR EMBL; AE00466; AAK6212.1; -.  
 DR InterPro; IPR001254; Ser\_protease\_Try.  
 DR Pfam; PF0089; Trypsin; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; UNKNOWN\_1.  
 KW Hydrolase; Serine protease.  
 SQ SEQUENCE 242 AA; 26583 MW; 6E72D8AEDEA2F1 CRC64;  
 Alignment Scores:  
 Pred. No.: 1.72e-67 Length: 242  
 Score: 988.00 Matches: 173  
 Percent Similarity: 91.71% Conservative: 15  
 Best Local Similarity: 84.39% Mismatches: 17  
 Query Match: 4 Indels: 0  
 Gaps: 0  
 US-09-598-982-20 (1-771) x Q96R25 (1-242)  
 QY 19 ATCGCGGGGGTCAGAGGCCAGGACACTGGCCCTGGCGAGGTGAGCCAGAGTC 78  
 DB 38 IleValGlyIgYIgIgIluIapProArgSerLysIrrPProRpGivInvalSerLeuArgVal 57  
 QY 79 CACGCCCATACTGGATGCACTTGCGGGGCCTCACCCACCCCCAGTGGTGCTG 138  
 DB 58 ArgGlyProtryTpMetHisPheCysGlyGlySerLeuLeHsProGlnTrpValLeu 77  
 QY 139 ACCGGCGCGCGCGTGGTGGACCGGACGTCAGGAGTCGGCCGCTCAGGGTCAACTG 198  
 DB 78 ThralalaAlaHisCysValGluProAspIleLySpLeuAlaAlaLeuArgValGlnLeu 97  
 QY 199 CGGGGACGACCTCTACTACCAAGGACCAAGCTGGCCGCTCACAGGATCATGGCAC 258  
 DB 98 ArgGluGlnHisLeuTyryrglnAspGlnLeuLeuProValSerArgIleIleValHs 117  
 QY 259 CCACGTTCTACACGCCGAGATGGAGGACATGGCCCTGCAGGAGCTGGAGGCC 318  
 DB 118 ProGlnPhytyrIleIleGlnThrGlyAlaAspIleAlaLeuLeuGluIleGluIleValHs 137  
 QY 319 GTGAAGGTCTCCAGGCCACGTCACCTGGCCCTGCCCCCTGAGACCTTCCC 378  
 DB 138 ValAsnIleSerSerHistIleIleThrValThrLeuProProAlaSerGluThrPhePro 157  
 QY 379 CCGGAGATGGCTGGGCACTGGGCACTGGCTGGGCACTGGACAATGAGGCTGAGAGTC 438  
 DB 158 ProGlyMetProCysIrrPvalIleGlyIgYIgIgIluAspValAspAsnValHsIlePro 177  
 QY 439 CCGCATTCTCTGAAAGGGTGAAGGCCATTAATGAAACACATTTGCCACCA 498  
 DB 178 ProProtryProLeuIgIluValGluValProValValGluAshnHsIleCysAsnAla 197

RESULT 7  
 Q9UQ18 PRELIMINARY; PRT; 166 AA.  
 ID Q9UQ18  
 AC Q9UQ18;  
 DT 01-MAY-2000 (TREMBrel. 13, last sequence update)  
 DT 01-MAR-2002 (TREMBrel. 20, last annotation update)  
 DE Mast cell MNCP-7-like I protein (fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OX NCBI\_TAXID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99121069; PubMed=9920877;  
 RA Palladino M., Reijo M.S., Shavesteh L., Blount J.L., Caughey G.H.;  
 RT "Characterization of genes encoding known and novel human mast cell tryptases on chromosome 16p13.3.";  
 RL J. Biol. Chem. 274:3355-3362(1999).  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.  
 DR HSSP; P2031; IMAO.  
 DR MEROPS; S01.054; -.  
 DR InterPro; IPR001314; Chymotrypsin.  
 DR PROSITE; PS00089; Trypsin; 1.  
 DR Pfam; PF00089; Trypsin; 1.  
 DR PRINTS; PRO072; CHMOTRPFIN.  
 DR SMART; SM00020; TRYV\_SPC; 1.  
 DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
 KW Hydrolase; Serine protease.  
 FT NON\_TER  
 SQ SEQUENCE 166 AA; 18433 MW; A7AE897A6423E7D5 CRC64;  
 Alignment Scores:  
 Pred. No.: 1.99e-45 Length: 166  
 Score: 699.00 Matches: 124  
 Percent Similarity: 94.85% Conservative: 5  
 Best Local Similarity: 91.18% Mismatches: 7  
 Query Match: 47.94% Indels: 0  
 DB: 4 Gaps: 0  
 US-09-598-982-20 (1-771) x Q9UQ18 (1-166)  
 QY 19 ATCGCGGGGGTCAGAGGCCAGGAGCTGGCCCTGGCAGGTGAGCTGAGAGTC 78  
 DB 31 IleValGlyIgYIgIgIluIapProArgSerLysIrrPProRpGivInvalSerLeuArgVal 50  
 QY 79 CACGCCCATACTGGATGCACTTGCGGGGCCTCACCCACCCCCAGTGGTGCTG 138  
 DB 51 ArgGlyProtryTpMetHisPheCysGlyGlySerLeuLeuLeHsProGlnTrpValLeu 70  
 QY 139 ACCGGCGCGCGTGGGCACTGGCTGGCCCTGAGGCTGAGGATCTGGCTGAGAGTC 198  
 DB 71 ThralalaAlaHisCysMetGluProAspIleLySpLeuAlaAlaLeuArgValGlnLeu 90  
 QY 199 CCGGAGACGACCTCTACTACCAAGGACCAAGCTGGCCGCTGAGGATCATGTC 258  
 DB 91 ArgGluGlnHisLeuTyryrglnAspGlnIleLeuProValSerArgIleIleValHs 110









Qy 592 GGGCAGCTCGGAGGCCCTGGTGCAAGGTGATGGACCTGGCTGAGGGCGCTG 651  
 ||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||  
 Db 247 GlyaspSerGly1g1yProLeuValCysAspSerLeasnAspValTrpIleAlaGlyVal 266  
 Qy 652 GTCAGCTGGGGCAGGGCTGGCACCCAGCCAAACGGCCCTGGCATCTACACCGTGCACC 711  
 ||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||  
 Db 267 ValserTrpIleSerAspSerAlaLeuPheLysArgProGlyAlaTyrrhRasnValser 286  
 ||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||  
 Qy 712 TACTCTGAGCTGATCCACCATTGCCCCAACGGCGTGAAGCAGGCCGC 765  
 ||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||  
 Db 287 ValtryIleSerTrpIleGlnAsnThrMetTrpAsnLeuProMetGluGlyArg 304  
 RESULT 15  
 Q9PVX7 PRELIMINARY; PRT; 389 AA.  
 ID Q9PVX7;  
 AC 09PVX7;  
 DT 01-MAY-2000 (TREMBLel. 13, Created)  
 DT 01-MAY-2000 (TREMBLel. 13, Last sequence update)  
 DE EPidermis specific serine protease.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
 OC Xenopodidae; Xenopus.  
 OX NCBI\_TAXON=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yamada K.;  
 RT "The expression control of xepsin by non-axial and planar posteriorizing signals in Xenopus epidermis";  
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 CC -! SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.  
 DR EMBL; AB018694; BAA84941.1; -  
 DR HSSP; P00763; IDPO;  
 DR InterPro; IPR01314; Chymotrypsin.  
 DR InterPro; IPR01254; Ser\_Protease\_Try.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR0072; CHYMOTRYPSIN.  
 DR SMART; SM00020; TRYF\_SPC\_1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
 DR PROSITE; PS00135; TRYPSIN\_SEB; 1.  
 KW Hydrolase; Protease; Serine Protease.  
 SQ SEQUENCE: 389 AA; 42375 MW; B31FB4A2FF5D1F6E3 CRC64;

Alignment Scores:

Pred. No.:	1.94e-33	Length:	389
Score:	543.00	Matches:	108
Percent Similarity:	57.14%	Conservative:	36
Best Local Similarity:	42.86%	Mismatches:	96
Query Match:	37.24%	Indels:	12
DB:	13	Gaps:	5

US-09-598-982-20 (1-771) x Q9PVX7 (1-389)

Qy 7 CTCGAGAAAAGAATGCTCGGGGTCAGGGCCCCAGGAGCAAGTGACCTGGCTGGAGGTG 66  
 ||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||  
 Db 22 IleSerAsnArgIleValGly1yMetAspSerIysArgGlyIutPrpItrpPglInle 41  
 Qy 67 AGCCTGAGAGTCACGCCATACTGGATGCCACTCTGGGGCTCCCTCATCCCC 126  
 ||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||  
 Db 42 SerLeuSerTrpIlySerAsp-----SerIleCysGlyIysSerLeuLeuThrAsp 58  
 Qy 127 CAGGGGCTGACGCCGGCTGGGGACGGAGTCAGGATCTGGCCCC 186  
 ||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||  
 Db 59 SerTrpValMetThrAlaLahIsCysIle-----AspSerLeuAspValSerTrpTyr 76  
 Qy 187 AGGGTCAACTGCGGAGCAGCTACTACAGGACAGCTGCTGCCGTAGCAG 246  
 ||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||  
 Db . 77 ThrValItyrLeuGlyYalatYrgInLeuSerAlaProAsn--SerThrValSerArg 95

Qy 247 -----ATCATCGTCACCCACAGTCTACACCCCGCATCCAGGCCGACTC 294  
 ||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||  
 Db 96 GlyvallySerSerIleThrIlyHisProAspPheIlyProAspIle 115  
 Qy 295 GCCCTGCTGGAGCTGGAGGAGCCGGTGAAGGCTCCAGCACCGTCACCGTC 354  
 ||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||  
 Db 116 AlaLeuIleGluIleGluIysProValIhrPheIhrProTyrlleProIleCysIle 135  
 Qy 355 CCCCTGCTAGACCTCCCCCGGGATGCCGCTGGCAGCAGGCTCCACCGTCACCGTC 414  
 ||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||  
 Db 136 ProSerGlnAspValGlnPheIalaAlaGlyIutIhrMetCysTrpValThrGlyIrpGlyAsn 155  
 Qy 415 GTGGACAATGATGAGGCCCTCCACCGCATTCTCTGAAAGGAGGAGGTC 474  
 ||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||  
 Db 156 IleGlnIglGluIlyIhrProLeuIleSerIysPheIleGlnIlyAlaGluValIala 175  
 Qy 475 ATGAAACACATTTGAGCAGCAAATACCCACCTGGGCCTACACGGAGACGTC 534  
 ||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||  
 Db 176 IleAspSerSerValCysGlyIutIhrMetIyrglUserSerIleGlyIutIleProAspIphe 195  
 Qy 535 CGCAVCGTCCGAGCAGCATGCTGGCCTACACGGAGACGACTC 588  
 ||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||  
 Db 196 SerPheIleGlnIlyAspMetValCysAlaGlyIutIlyArgIleAspAlaCys 215  
 Qy 589 CAGGCAGCTCCGGAGGCCCTGGTGCAGGCTGAATGCCACCTGGCTGCAAGGG 648  
 ||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||  
 Db 216 GlyIyAspSerGlyIlyProLeuValCysAlaValAsnValAsnValIleAspIle 235  
 Qy 649 GTGGCAGCTGGGGAGGCCCTGGTGCAGGCTGAATGCCACCTGGCTGCA 708  
 ||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||  
 Db 236 IleValSerTrpIgIyTyrglyCysAlaGluProAsnArgProGlyIutIhrLysVal 255  
 Qy 709 ACCRACTACTGGACTGGATCCACCACTATGTC 744  
 ||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||  
 Db 256 GlyIutIlyArgIleAspIleIlyProLeuIleAspIleAspIleAspIle 267

Search completed: January 31, 2003, 06:58:26  
 Job time : 63.5 secs





AC	Q15661; Q15663; Q9H2Y4;
DT	16-OCT-2001 (Rel. 40, Created)
DT	16-OCT-2001 (Rel. 40, Last sequence update)
DE	TRYPTASE beta-1 precursor (EC 3.4.21.59) (Tryptase 1) (Tryptase I).
GN	TSPB1.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TAXID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=90251647; PubMed=2187193;
RA	Vanderlinde P., Ballinger S.M., Tam E.K., Goldstein S.M., Craik C.S.,
RA	Caughey G.H.;
RT	"Characterization of genes encoding known and novel human mast cell tryptases on chromosome 16p13.3."
RL	J. BIOL. CHEM. 274:3355-3362(1999).
RN	[2]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=99121069; PubMed=9920877;
RA	Pallairo M., Feijoo M.S., Shayesteh L., Blount J.L., Caughey G.H.,
RA	Wang H.W., McNeil H.P., Thomas P.S., Murphy B.N., Webster M.J.,
RA	Bettaratchai A., King G., Heywood G.J., Huang C., Stevens R.L.,
RA	Hunt J.E.;
RT	"Molecular cloning and characterization of novel human trypsin and splicing variants."
RT	Submitted (NOV-1999) to the EMBL/GenBank/NCBI databases.
-i	FUNCTION: TRYPTASE IS THE MAJOR NEUTRAL PROTEASE PRESENT IN MAST CELLS AND IS SECRETED UPON THE COUPLED ACTIVATION-DEGRANULATION RESPONSE OF THIS CELL TYPE.
-i	CATALYTIC ACTIVITY: preferential cleavage: Arg- -, Lys- -, but with more restricted specificity than trypsin.
-i	SUBCELLULAR LOCATION: RELEASED FROM THE SECRETORY GRANULES UPON MAST CELL ACTIVATION.
-i	ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are produced by alternative splicing.
-i	SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. TRYPTASE SUBFAMILY.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to licensee@isb-sib.ch).
DR	EMBL: M3394; AAC83172; -.
DR	EMBL: M3391; AAA36778; 1; -.
DR	EMBL: AF099144; AAD17860; 1; -.
DR	EMBL: AF206667; AAG35697; 1; -.
DR	HSSP: P20231; JAOL.
DR	MEROPS: S01_242; -.
DR	GeneBank: HGNC:12019; TPSB1.
DR	MM: 191081; -.
DR	InterPro: IPR01314; Chymotrypsin.
DR	InterPro: IPR01254; Ser-protease_Try.
DR	Pfam: PF00089; trypsin; 1.
PRINTS:	PRO0722; CHYMOTRYPSIN.
DR	SMART: SM00020; TRY_SP_C; 1.
DR	PROSITE: FS00240; TRPSIN_DOMAIN; 1.
DR	PROSITE: PS00134; TRPSIN_HIS; 1.
DR	PROSITE: PS0015; TRPSIN_SER; 1.
KW	Hydrolyase; Serine protease; Signal; Glycoprotein; Zymogen; Alternative splicing.
FT	SIGNAL 1 18 POTENTIAL.
PROPEP	19 30 ACTIVATION PEPTIDE (BY SIMILARITY).
FT	CHAIN 31 275 TRYPTASE_BETA-1.
FT	ACT_SITE 74 74 CHARGE_RELAY_SYSTEM.
FT	ACT_SITE 121 121 CHARGE_RELAY_SYSTEM.
FT	ACT_SITE 224 224 CHARGE_RELAY_SYSTEM.
FT	DISULFID 59 75 BY_SIMILARITY.
FT	DISULFID 155 230 BY_SIMILARITY.
FT	DISULFID 188 211 BY_SIMILARITY.
FT	DISULFID 220 248 BY_SIMILARITY.
FT	CARBONYD 132 132 MISSING (IN ISOFORM 2).
FT	CARBONYD 233 233 N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	VARSLIC 79 87 MISSING (IN ISOFORM 2).
SQ	SEQUENCE 275 AA; 30515 MW; ADC48FDC51F37112 CRC64;
DB:	Alignment Scores: Pred. No.: 1.38e-72 Score: 1363.00 Percent Similarity: 99.18% Best Local Similarity: 99.18% Query Match: 93.48% DB: 1 Gaps: 0
US-09-598-982-20 (1-771) x TRB1_HUMAN (1-275)	QY 19 ATCGTCGGGGTCAGGAGGCCAGGAGCAAGTGCCCTGGCAGGTGAGCTGAGATC 31 IleValGlyIleGlyIleGluIleAlaProArgSerLeuIleProrPrGlnValSerLeuIargVal Db 79 CACGCCCTACTGGACTCTGGCGGGCCRCCTCATCCACCCCAGTGGCCGCG 71 ThralaAlaHisCysValIleGlyProAspValLysPheIleAlaAlaLeuIrgValGlnLeu Db 51 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnIntrValLeuI QY 199 CGGGAGCACCTCTACTAACGAGCACAGCTGGCAGGAGCTGTCGGCTGAGCG 91 ArgGluGlnHistLeuTyRyGlnIaspGlnIleLeuIleProValSerArgIleIleIleValHis Db 259 CCAGCTCTACAGGCCAGAGCAGCGGACTCTGCCTGCGAGCTGGAGGAGCG 111 ProGlnIphErythrAlaGlnIleGlyIalaAspIleAlaIleLeuIleGluIleGluIle Db 319 GTGAGGGCTCAGCCACGCTCCACAGGTACCGTCCCCCTGCCTCAGAGCTTCCCC 131 ValAsnValSerSerHisValHistValThrIleProProAlaSerGluIlePhePrc Db 379 CCGGGATGCCGTCTGGTCACTGCGTGGGGCATGTGGACAATGATAGCCTCCCA 151 ProGlyMetProCysTrpIleValGlyTrpGlyIaspValIaspAsnAspGluIlePro Db 439 CCCGCAATTCTCGAAGCAGGAGGAAGGAACTATGGAAACACAATTGGAGCCCA 171 ProPropheProLeuLySgInvalIlySvalProIleMetGluAsnHisIleCysAspAla Db 499 ARATTCACCTGGCCCTACACGGAGACGACCTCCGCATCGCTGGTACGAGCTGCTG 619 Ile QY 559 TGGCCGGAACACCCGGAGGACTCATGCCAGGGACTCCGGGACTCCGGCCACCACT 211 CysAlaGlyIasnThrArgArgAspSerCysGindlyAspSerGlyGlyProLeuValCys Db 191 LystyRHisLeuGlyIalAtyRthrGlyAspAspValArgIleValArgAspAspMetIeu Db 191 LystyRHisLeuGlyIalAtyRthrGlyAspAspValArgIleValArgAspAspMetIeu QY 679 CCCAACGGCTGCATCACCCGGTCACTACTGGACTGGATCCACCACT 231 LysValAsnGlyIleTrpLeuGlnIalGlyIvaValSerTrpGlyIgluIycysAlaGin QY 739 GTCGCGAAAGCCG 753

us-09-598-982-20.n2p.esp

RESULT	3	DB	271	ValProIysLysPro	275
TRYA_HUMAN		ID	TRYA_HUMAN	STANDARD;	PRT; 275 AA.
AC	P15157; Q9UQ11; O9H2Y5;	DT	01-APR-1990 (Rel. 14, Created)		
	16-OCT-2001 (Rel. 40, last sequence update)	DE	Alpha-trypsinase precursor (EC 3.4.21.59) (Tryptase 1).		
CC	Eukaryota; Metazoa; Chordata; Craniata; vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. sapiens (Human).	OC			
CC	NCB1_1 TAXID=9606;	OX			
RN	[1] SEQUENCE FROM N.A. (ISOFORM 1).	DR			
RC	TISSUE=Lung;	DR			
RX	MEDLINE=9000311; PubMed=267049;	DR			
RA	MILLER J.S., WESTIN E.H., SCHWARTZ L.B.;	DR			
RT	"Cloning and characterization of complementary DNA for human trypsinase.";	DR			
RL	J. Clin. Invest. 84:1188-1195(1989).	DR			
RN	[2] REVISIONS TO 89-93 AND 108.	DR			
RP	Schwartz L.B.; Submitted (MAR-1990) to the EMBL/GenBank/DDBJ databases.	DR			
RL	[3] SEQUENCE FROM N.A. (ISOFORM 1).	DR			
RP	MEDLINE=99121069; PubMed=9920877;	DR			
RX	Palmaoro M., Feijo M.S., Shareefeh L., Blount J.L., Caughey G.H.;	DR			
RT	"Characterization of genes encoding known and novel human mast cell tryptases on chromosome 16p13.3."	DR			
RL	J. Biol. Chem. 274:3355-3362(1999).	DR			
RN	[4] SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).	DR			
RP	PISSUE=Lung; Wang H.W., McNeil H.P., Thomas P.S., Murphy B.N., Webster M.J., Hettiaratchi A., King G., Heywood G.J., Huang C., Stevens R.L., Hunt J.E.;	DR			
RT	"Molecular cloning and characterization of novel human trypsin and splicing variants."	DR			
RL	Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.	DR			
RN	[5] SEQUENCE OF 31-50, AND PITUITARY, SEQUENCE OF 31-38.	DR			
RT	TISSUE=Lung;	DR			
RC	MEDLINE=87109258; PubMed=3543004;	DR			
RX	Cronish J.A., Seidah N.G., Marcinkiewicz M., Hamelin J., Johnson D.A., Chreistin M.;	DR			
RA	"Human pituitary trypsinase: molecular forms, NH2-terminal sequence, immunocytochemical localization, and specificity with prohormone and fluorogenic substrates."	DR			
RL	J. Biol. Chem. 262:1363-1373(1987).	DR			
CC	-!- FUNCTION: TRYPTASE IS THE MAJOR NEUTRAL PROTEASE PRESENT IN MAST CELLS AND IS SECRETED UPON THE COUPLED ACTIVATION-DEGRANULATION RESPONSE OF THIS CELL TYPE.	DR			
CC	-!- CATALYTIC ACTIVITY: Preferential cleavage: Arg- -, Lys- -, but with more restricted specificity than trypsin.	DR			
CC	-!- SUBUNIT: HOMOTRIMER.	DR			
CC	-!- SUBCELLULAR LOCATION: RELEASED FROM THE SECRETORY GRANULES UPON MAST CELL ACTIVATION.	DR			
CC	-!- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are produced by alternative splicing.	DR			
CC	-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. TRYPTASE SUBFAMILY.	DR			
CC	-----	DR			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the European Bioinformatics Institute and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).	DR			

RL PROC. Natl. Acad. Sci. U.S.A. 87:3230-3234(1990).  
 QY 379 CCGGGATCCGAGCTGGCTCACTGGCTGGGGATGTGGACATGAGGCCCTCCA 438  
 CC -1- FUNCTION: TRYPSINE IS THE MAJOR NEUTRAL PROTEASE PRESENT IN MAST  
 Db 151 ProgluMetProCysTrpValThrLysPheGlyAspValAsnAspGluLeuPro 170  
 CC CELLS, AND IS SECRETED UPON THE COUPLED ACTIVATION-DGRANULATION  
 CC RESPONSE OF THIS CELL TYPE.  
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-1-, Lys-1-, but  
 CC with more restricted specificity than trypsin.  
 CC C-terminus: C-terminal peptide: C-terminal peptide: C-terminal peptide:  
 CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; a long form (shown here) and a  
 CC short form; are produced by alternative splicing. The short form  
 CC is probably a non functional variant.  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. TRYPSINE SUBFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).  
 CC  
 DR EMBL: M57626; AAA39988; 1;  
 DR EMBL: M77625; AAA39987; 1;  
 DR EMBL: X18183; AA339725; 1;  
 DR EMBL: X178542; CA55288; 1;  
 DR PIR: A38654; A38654.  
 DR PIR: D35646; D35646.  
 DR HSSP; P20231; IAO.  
 DR MEROPS; S01.025; -.  
 DR MGD; MGI:96942; Mpt6.  
 DR InterPro; IPR01314; Chymotrypsin.  
 DR InterPro; IPR01254; Serine\_protease\_Try.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHIMOPRPSIN.  
 DR SMART; SM00020; TRYPSIN\_SPC; 1.  
 DR PROSITE; PS50240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Hydrolase; Serine protease; Signal; Glycoprotein; Zymogen;  
 KW Alternative Splicing.  
 FT SIGNAL 1  
 FT PROPEP 22  
 FT CHAIN 32  
 FT ACT\_SITE 75  
 FT ACT\_SITE 75  
 FT ACT\_SITE 122  
 FT ACT\_SITE 225  
 FT DISULFD 60  
 FT DISULFD 156  
 FT DISULFD 189  
 FT DISULFD 221  
 FT CARBOHYD 133  
 FT VANSPLIC 230  
 FT VANSPLIC 231  
 SQ SEQUENCE FROM N.A.;  
 RC STRAIN=leader X A1;  
 RX MEDLINE-94032807; PubMed=8210998;  
 RA Nilsson R., Abrikosoff H., Gobl A.E., Nilsson G., Aveskog M., Larsson L.G.,  
 RT "Expression of a mast cell tryptase in the human monocytic cell lines  
 RT U-373 and Mono Mac 6.",  
 RL Scand. J. Immunol. 38:359-367(1993).  
 RN [3]  
 SEQUENCE FROM N.A.;  
 RP STRAIN=leader X A1;  
 RC MEDLINE-9504852; PubMed=759952;  
 RA Huang R., Hellman L.T.;  
 RT "Genes for mast-cell serine protease and their molecular evolution.",  
 RL Immunogenetics 40:397-414(1994).  
 RN [4]  
 SEQUENCE OF 32-54.  
 RP MEDLINE-90022202; PubMed=2326280;  
 RA Reynolds D.S., Stevens R.L., Lane W.S., Carr M.H., Austen K.F.,  
 RA Serafini W.E.;  
 RT "Different mouse mast cell populations express various combinations  
 of at least six distinct mast cell serine proteases.",  
 US-09-598-982-20 (1-771) x MCT6\_MOUSE (1-276)  
 QY 4 CCGCTCGAGAAAGA-----AVGTCGGGGTCAGGGCCCCAGGAGAAGTGCCC 57  
 DB 25 ProAlaAsnGlnArgValGlyIleValGlyLysHisGluAsnSerGluSerLysTrpPro 44  
 QY 58 TGCAGCTGAGCTGAGTCACGCCCATCTGGATGCCCTCGGGGCTCCTC 117  
 DB 45 TrpGlnValSerLeuArgPheLysLeuAsnTryTrpIleHisPheCysLysGlySerLeu 64  
 QY 118 ATGCCACCCCACTGGGCTGAGCCGGGGTGGGGAGGAGTCAGGATCTG 177

Db 65 IleHisProGlnTrpValLeuThrAlaLahisCystAlaGlyProHisIleSerPro 84  
 QY |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
 178 GCGCCCTCAGGGTCACTGCGGGAGCAGCACCTACTACAGGCCAGATGGAGGCCACAGTC 237  
 CC |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
 Db 85 GlnLeuPheArgValGlnLeuArgLugIntrLeuTyrrTyrglyAspGlnLeuSer 104  
 QY |||||||:|||||:|||||:|||||:|||||:|||||:|||||:  
 238 CTGCTGAGCATCTGACGCCAACAGCTCACGCCAGATGGAGGCCACAGTCACCGTAC 297  
 CC |||||||:|||||:|||||:|||||:  
 Db 105 LeuAsnArgIleValHisProHistTyrrTyrrThraLugIlyAlaAspValAla 124  
 QY |||||||:|||||:  
 298 CTGCTGAGCTGGAGGAGCCGGGAAGGCTCCAGGCCACAGTCACCGTACCCGCC 357  
 CC |||||||:  
 Db 125 LeuIleGluLeuGluValProValAsnValSerThrHisIleHisProLeuSerLeuPro 144  
 DR :::::|||||:|||||:|||||:  
 DR PIR; A34410; A32410.  
 DR HSSP; P20231; IAAO.  
 DR MEROPS; S01.143;  
 DR InterPro; IPR01314; Chymotrypsin.  
 DR InterPro; IPR01254; Ser\_protease\_Try.  
 DR Pfam; PF00089; trypsin\_1.  
 DR PRINTS; PR00722; CHYMOTRYPTIN.  
 DR PROSITE; SM0050; TRYPSIN\_SPC; 1.  
 DR PROSITE; PS050240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 DR Hydrolase; Serine protease; Signal; Glycoprotein; Zymogen.  
 FT SIGNAL 1 20 POTENTIAL.  
 FT PROPEP 21 30 ACTIVATION PEPTIDE (BY SIMILARITY).  
 FT CHAIN 31 275 TRYPTASE.  
 FT ACT\_SITE 74 74 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 121 121 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 224 224 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT DISULFID 59 75 BY SIMILARITY.  
 FT DISULFID 155 230 BY SIMILARITY.  
 FT DISULFID 188 211 BY SIMILARITY.  
 FT DISULFID 220 248 BY SIMILARITY.  
 FT CARBOHYD 132 132 N-LINKED (GLICNAc.  
 SQ SEQUENCE 275 AA; 30088 MW; C3BB65251F248D5B CRC64;  
 Alignment scores:  
 Pred No : 6.13e-56 Length: 275  
 Score: 107.00 Matches: 189  
 Percent Similarity: 85.71% Conservative: 21  
 Best Local Similarity: 77.14% Mismatches: 35  
 Query Match: 73.87% Deletions: 0  
 DB: 1 Gaps: 0

RESULT 5

TRYT\_CANFA STANDARD; PRT; 275 AA.

ID TRYT\_CANFA

AC P15944;

DT 01-APR-1990 (Rel. 14, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Tryptase precursor (EC 3.4.21.59).

OS Canis familiaris (Dog).

OC Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Buteleostomi; OX NCBI\_TaxId=9615;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-8935460; PubMed=2504277;

RA Vanderslice P., Craik C.S., Nadel J.A., Caughey G.H.;

RT "Molecular cloning of dog mast cell tryptase and a related protease: structural evidence of a unique mode of serine protease activation";

RL Biochemistry 28:4148-4155(1989).

CC !- FUNCTION: TRYPTASE IS THE MAJOR NEUTRAL PROTEASE PRESENT IN MAST CELLS AND IS SECRETED UPON THE COUPLED ACTIVATION-DEGRANULATION RESPONSE OF THIS CELL TYPE.

CC !- CATALYTIC ACTIVITY: Preferential cleavage: Arg- -, Lys- , but with more restricted specificity than trypsin.

CC !- SUBUNIT: HOMOTRIMER.

CC !- SUBCELLULAR LOCATION: RELEASED FROM THE SECRETORY GRANULES UPON MAST CELL ACTIVATION.

CC !- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. TRYPTASE SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage for commercial entities requires a license agreement (See <http://www.ish-sib.ch/announce/> or send an email to license@ish-sib.ch).

CC EMBL; M24664; AAH30854.1; -.

CC PIR; A34410; A32410.

CC HSSP; P20231; IAAO.

CC MEROPS; S01.143;

CC InterPro; IPR01314; Chymotrypsin.

CC InterPro; IPR01254; Ser\_protease\_Try.

CC Pfam; PF00089; trypsin\_1.

CC PRINTS; PR00722; CHYMOTRYPTIN.

CC PROSITE; SM0050; TRYPSIN\_SPC; 1.

CC PROSITE; PS050240; TRYPSIN\_DOM; 1.

CC PROSITE; PS00134; TRYPSIN\_HIS; 1.

CC PROSITE; PS00135; TRYPSIN\_SER; 1.

CC Hydrolase; Serine protease; Signal; Glycoprotein; Zymogen.

FT SIGNAL 1 20 POTENTIAL.

FT PROPEP 21 30 ACTIVATION PEPTIDE (BY SIMILARITY).

FT CHAIN 31 275 TRYPTASE.

FT ACT\_SITE 74 74 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT ACT\_SITE 121 121 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT ACT\_SITE 224 224 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT DISULFID 59 75 BY SIMILARITY.

FT DISULFID 155 230 BY SIMILARITY.

FT DISULFID 188 211 BY SIMILARITY.

FT DISULFID 220 248 BY SIMILARITY.

FT CARBOHYD 132 132 N-LINKED (GLICNAc.

SQ SEQUENCE 275 AA; 30088 MW; C3BB65251F248D5B CRC64;

Alignment scores:

Pred No : 6.13e-56 Length: 275

Score: 107.00 Matches: 189

Percent Similarity: 85.71% Conservative: 21

Best Local Similarity: 77.14% Mismatches: 35

Query Match: 73.87% Deletions: 0

DB: 1 Gaps: 0

US-09-598-982-20 (1-771) x TRYT\_CANFA (1-275)

QY 19 ArgArgCggGGGTCAGGAGGCCAGGAGCAAGTGCCCTGCAGGAGGAGCTGAGTC 78

Db 31 IleValIgIlyArgLugIalaproGlySerlyStrpproTrpDlnValSerIleArgIle 50

QY 79 CAGGCCCATAGTGTGACTCTGCGGGGCTCCATGCCACCCAGGGCTGGCT 138

Db 51 LysGlyGlnTyrrTrpArgHistileCysGlyGlySerIleUllieHisProGlnTrpValLeu 70

QY 139 ACCCCGGGGCGTGGCGGACGGACGCTCAAGATCAGGCCCTCAGGGTGCAACTG 98

Db 71 ThrlAlaAlaLahisCysValGlyProAsnValCysProGluIleArgValGlnLeu 90

QY 199 CGGGAGCAGCACCTCTACTACAGGACAGCAGCTGGCGGAGCAGAGATCGGCCAC 258

Db 91 ArgGluGlnHisLeuTyrrTyrglyAlaAspIleLeuLeuProValAsnArgIleValMetHis 110

QY 259 CCACAGTCTACCGCCAGATCGGAGCCAGTCAGGCCCTGCCCCCTGGAGCTGGAGGCC 318

Db 111 ProAsnTyrrTyrrProGluIleGluLeuLeuProValAsnArgIleValMetHis 130

QY 319 GTGAGGCTCCAGCCAGTCCACGCTACACCTGCCTGCCCCCTGGAGCTGGAGGCC 378

Db 131 ValAsnValSerAlaLahisCysValGlnProValThrLeuProProAlaLeuGlnThrPhePro 150

QY 379 CGGGAGGCGGCGTGGCTACTGGCGGCGAGTGGAGCATGATGAGCGGCCCTCCA 438

Db 151 ThrglylhrProCysTrpValIgIlyTrpGlyAspValHisSerGlyLysValProLeuPro 170

QY 439 CGCCATTCTCTGAAGCAGSTGAAGTCCCATAATGAAACCACATTGTGACCA 498

Db 171 ProProPheProLeuLysGlnValValProLeuValGluAsnSerMetCysAspVal 190

QY 499 AATACCACCTGGCCCTACACGGAGACGACTCGCATCGCCTGACGAGCTG 558  
 ::|||||::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Db 191 GlnTyrHisLeuGlyLeuSerThrGlyAspGlyAlaArgLysAspMetLeu 210  
 ::|||||::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 QY 559 TGTGCCGGAACACCGGGAGGGACTCATGCCAGGGCAGCTGGGGCCCTGTGTC 618  
 ::|||||::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Db 211 CysAlaGlyAsnSerLySsSerAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 230  
 ::|||||::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 QY 619 AAGGGTAATGCGACCTGGCGACGGGACTCATGCCAGGGCAGCTGGGGCCCTGTGTC 678  
 ::|||||::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Db 231 ArgValArgLyValTrpLeuGlnAlaGlyValSerTrpGlyGluGlyCysAlaGln 250  
 ::|||||::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 QY 679 CCCACCGGCCTGGCATCACCGGTGACCTACTACTTGAGTTGGATCACCCTAT 738  
 ::|||||::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Db 251 ProAsnArgProGlyLysTyrThrArgValAlaTyrTyrLeuAspTrpLeuAsp 270  
 ::|||||::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 QY 739 GTCCCCAAAGGCC 753  
 ::|||||::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Db 271 ValProLySgluPro 275  
 ::|||||::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 RESULT 6  
 TRYT\_MERUN STANDARD PRT: 270 AA.  
 ID TRYT\_MERUN PRT: 270 AA.  
 AC P50342;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DE 16-OCT-2001 (Rel. 40, Last annotation update)  
 Mast cell tryptase precursor (BC 3.4.21.59).  
 Meriones unguiculatus (Mongolian Jird) (Mongolian gerbil).  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Gerbillinae;  
 Meriones; [1]  
 RN [1]  
 RP STRAIN=MG/SEA; TISSUE=Intestine;  
 RX MEDLINE=9356971; PUBMED=7639711;  
 RA Murakumo Y., Ide H., Itoh H., Tomita M., Kobayashi T.,  
 RT Maruyama H., Horii Y., Nawa Y.;  
 RT "Cloning of the cDNA encoding mast cell tryptase of Mongolian gerbil,  
 RT Meriones unguiculatus, and its preferential expression in the  
 RT intestinal mucosa.";  
 RL Biochem. J. 309:921-926(1995)  
 CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-!, Lys-!, but  
 CC with more restricted specificity than trypsin.  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY SI. TRYPTASE SUBFAMILY.  
 CC  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to license@isb-sib.ch).  
 CC  
 EMBL; D31789; BRA06598.1; -.  
 DR pfam: PF0089; trypsin; 1.  
 DR HSSP; P20331; IAAO.  
 DR MEROPS; S01.143; -.  
 DR InterPro; IPR00134; Chymotrypsin.  
 DR InterPro; IPR001254; Ser\_protease\_Try.  
 DR Prints; PR00722; Chymotrypsin.  
 DR SMART; S000020; Tryp\_Spc; 1.  
 DR PROSITE; P55020; TRYPIN\_DOM; 1.  
 DR PROSITE; P50014; TRYPIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPIN\_SER; 1.  
 KW Hydrolase; Serine protease; signal; Glycoprotein.  
 FT SIGNAL 1 25 POTENTIAL.  
 FT CHAIN 26 270 MAST CELL TRYPTASE.  
 FT ACT\_SITE 69 69 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 116 116 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 219 219 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT  
 FT DISULFID 54 70 BY SIMILARITY.  
 FT DISULFID 150 225 BY SIMILARITY.  
 FT DISULFID 183 206 BY SIMILARITY.  
 FT DISULFID 215 243 BY SIMILARITY.  
 FT CARBONYD 46 46 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBONYD 127 127 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 228 228 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 270 AA; 30166 MW; IBEI02DB86943401 CRC64;  
 DB:  
 Alignment Scores:  
 Pred. No.: 8.02e-56 Length: 270  
 Score: 1075.00 Matches: 188  
 Percent Similarity: 84.77% Conservative: 18  
 Best Local Similarity: 77.37% Mismatches: 37  
 Query Match: 73.73% Indexes: 0  
 DB:  
 Gaps: 0  
 US-09-598-982-20 (1-771) × TRYT\_MERUN (1-270)  
 QY 19 ATCGTGGGGTCAGGAGCCCAGGAGCACTGGCCCTGCGAGGTGACCTGAGACTC 78  
 Db 26 ILevalGlyGlyGlyGluAlaProGlyAlaLysTrpProTrpGlnValSerLeuArgAla 45  
 QY 79 CACGGCCCATPACTGATGCACTTCGCGGCGCTCCATCANNACCCCCAGTGGTGCNG 138  
 ::|||||::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Db 46 AsnGluThrTyrTrpArgHsPheCysGlyGlySerLeutLeuHsProGlnTrpValLeu 65  
 QY 139 ACCGGCGGCGGTGGTGGACGACGTCAGGATCTGGCGCTCAGGGTCAACTG 198  
 ::|||||::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Db 66 ThraAlaAlaLysIleValGlyProHsAlaAspProAsnLysValArgValGlnLeu 85  
 QY 259 CACAGTTCAACGCCAGATCGAGCGACATGCCCTGCTGAGCTGGAGGAGGAGC 318  
 ::|||||::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Db 106 ProTrpPhenylTyrAlaThrGlnAsnGlyLysAspIleAlaLeuLeuLysAsnPro 125  
 QY 319 CTGAAGGTCTCACGCCACGTCACAGGACCCCTGCCCTGCTCACAGACCTCCC 378  
 ::|||||::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Db 126 ValAsnLeuSerSerHsValHistroValLeuLeuLysAsnPro 145  
 QY 379 CGGGGATGCCGCTGGTCACTGGCTGGCATGGCAATGAGCCCTCCA 438  
 ::|||||::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Db 146 SerGlyThrIleCysTrpValThrGlyTrpGlyAsnIleAspAspValSerLeuPro 165  
 QY 439 CGCCCATTTCTCTGAAGCAGGTGGTCCCTAATGAAACCACATTGTGACGCA 498  
 ::|||||::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Db 166 ProProPheProLeuLysGluValGlnValProValValGluAsnGlnLeuLeuAspLeu 185  
 QY 499 AATACCACCTGGGCCATCACACGGAGACGAGCAGCAGCAGCTGGCGATCGCCTGACGATG 558  
 ::|||||::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Db 186 DystyRHisTyrGlyValTrpThrGlyAspAsnIleHistLeuValArgAspAspLeu 205  
 QY 559 TGTGCCGGACACCGGGAGGGACTCATGCCAGGGACTCCGGAGGCCCTGGTGC 618  
 ::|||||::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Db 206 CysAlaGlyLysIleValSerCysGlnGlyAspSerGlyGlyProLeuValCys 225  
 QY 619 AGGTCATGCACTGGCAGCAGGGGGGGTCAGCTGGGGAGGCCTGGGCCAG 678  
 ::|||||::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Db 226 LysValAsnGlyThrTripleGlnAlaGlyValSerTrpGlyGluGlyProLeuVal 245  
 QY 679 CCCACCGGCCGCTGGCATCTAACCGGTGACCTACTACTTGAGCTGGTCACACT 738  
 ::|||||::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Db 246 ProAsnArgProGlyLysTyrThrArgValThrTyrTyrLeuAspTrpLeuAspArgTyr 265  
 QY 739 GTCCCCAA 747  
 ::|||||:  
 Db 266 ValProLys 268  
 RESULT 7  
 MCT7\_MOUSE

ID MCT7\_MOUSE STANDARD: PRT; 273 AA.  
AC Q02844;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Mast cell protease 7 precursor (EC 3.4.21.59) (MNCP-7) (Trypsinase).  
GN MCF7.  
OS MUS musculus (Mouse)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TAXID=10090;  
RN [1]  
RP SEQUENCE FROM N.A. (TRUNCATED ISOFORM).  
RC STRAIN=C57BL/6J;  
RX MEDLINE=93087489; PubMed=1454796;  
RA McNeil H., Reynolds D.S., Schiller V., Ghildyal N., Gurley D.S.,  
RA Austen K.F., Stevens R.L.;  
RT "Isolation," characterization, and transcription of the gene encoding  
mouse mast cell protease 7.;  
RL Proc. Natl. Acad. Sci. U.S.A. 89:11174-11178(1992).  
RN [2]  
RP SEQUENCE FROM N.A. (TRUNCATED ISOFORM).  
RC STRAIN=C57BL/6J;  
RA Hunt J.E., Stevens R.L., Austen K.F., Zhang J., Xia Z.,  
RA Ghildyal N.;  
RT "Natural disruption of the mouse mast cell protease 7 gene in the  
C57BL/6 mouse";  
RL J. Biol. Chem. 271:2851-2855(1996).  
CC -!- FUNCTION: TRYPSINASE IS THE MAJOR NEUTRAL PROTEASE PRESENT IN MAST CELLS AND IS SECRETED UPON THE COUPLED ACTIVATION-DEGRANULATION RESPONSE OF THIS CELL TYPE.  
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-, Lys-|-, but with more restricted specificity than trypsin.  
CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; a long form (shown here) and a truncated form; are produced by alternative splicing. The alternative splicing event is due to a G to A point mutation at the exon 2/intron 2 splice site and causes loss of protein expression. The alternatively spliced transcript is only found in C57BL/6 mouse.  
CC -!- DEVELOPMENTAL STAGE: IS NOT EXPRESSED IN MATURE SEROSAL OR MUCOSAL MAST CELLS AND IS EXPRESSED ONLY TRANSIENTLY AT AN EARLY STAGE OF IN VITRO MAST CELL DIFFERENTIATION.  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. TRYPSINASE SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

MGB: MG:1:96943; Mcp7.  
EMBL: L0053; AAA39592.1; -.  
DR Interpro; IPR001314; Chymotrypsin.  
DR Interpro; IPR001254; Ser\_Protease\_Try.  
PFAM; PF00089; trypsin; 1.  
PRINTS; PR00722; CHYMOTRYPSIN.  
SMART; SM00020; TRY\_SPC; 1.  
MEROPS; S01\_026; -.  
HSSP: P20331; IAAO.  
DR M00000; S01\_026; -.  
DR MGD: MG:1:96943; Mcp7.  
DR Interpro; IPR001314; Chymotrypsin.  
DR Interpro; IPR001254; Ser\_Protease\_Try.  
PFAM; PF00089; trypsin; 1.  
PRINTS; PR00722; CHYMOTRYPSIN.  
SMART; SM00020; TRY\_SPC; 1.  
PROSITE; PS00134; TRYPSEN\_HIS; 1.  
PROSITE; PS00135; TRYPSEN\_SER; 1.  
Hydrolase; Serine protease; Signal; Zymogen; Alternative splicing; Glycoprotein. 1 18 POTENTIAL.  
FT PROPEP 19 28 ACTIVATION PEPTIDE (POTENTIAL).

Alignment scores:  
Pred. No.: 8.95e-55 Length: 273  
Score: 1057.00 Matches: 189  
Percent Similarity: 82.07% Conservative: 17  
Best Local Similarity: 75.30% Mismatches: 43  
Query Match: 72.50% Indels: 2  
DB: 1 Gaps: 1

US-09-598-982-20 (1-771) x MCT7\_MOUSE (1-273)

OY 1 GGGCCCCCTGGAGAAAAGA-----ATCGCGGGGTCAAGGAGCCCCCAGGAGCAAGTGG 54  
Db 21 GLYProAlaMetThrArgLysGluGlyIleValGlyGlyGlyGlyGlyInGluLahIslyAsnLysTrp 40  
QY 55 CCCPGGCAGGTGGCCCTGAGAGTCACGCCCATACTGTGACTCTGCGGGG6STCC 114  
Db 41 ProTrpGlnValSerLeuArgIlaAsnAspThrTyrTrpMetHisPheCysGlyGlySer 60  
Db 115 CTCAATCACCCTGGCTGACGCCGCCGGCGCCTGGTGACGGGACCCAAAGAT 174  
Db 61 LeuIleHisProGlnIleTrpValLeuThrAlaAlaHisCysValGlyProAspValAlaAsp 80  
QY 175 CTTGGCCGCTTCAAGGGGACCTACGGGAGGAGCACCTACCCAGGACCTAC 234  
Db 81 ProAsnLysValArgIleValGlyGlyGlyGlyIleTrpTyrTrpIleAspHisLeuMet 100  
QY 235 CGGGCAGGAGAACATCGTGACCCAGCTCTACGCCAGATGGAGGGACATC 294  
Db 101 ThrValSerGlnIleIleThrHisProAspPheTyrIleValGlnAspGlyValAspIle 120  
Db 295 GCGCTGCTGAGCTGGAGGAGCCGGTGGAGGTCTCCACGCCAACGTCACCTG 354  
Db 121 AlaLeuLeuLysLeuThrAsnProValAsnLeuSerAspPheValIleValHisProValProIle 140  
QY 355 CCCCTGGCTCAGAGACCCTCCCCGGATGCCGCTGGTACGGTACCGCTGGGCGAT 414  
Db 141 ProProAlaSerGluThrPheProSerGlyLysIleValGlyIleValAsn 160  
QY 415 GTGGACATGATGAGGCCCTCCCACCGCCATTCTCTGAGGAGGTGAGGCCATA 474  
Db 161 IleAspAsnGlyIleValAsnLeuProProProPheProLeuIleValGlnValProIle 180  
QY 475 ATGGAAACACCATGTTGAGCCTAACATGCCACCTGGCGCTACACGGAGGACCTC 534  
Db 181 IleGluAsnHisLeuCysAspLeuLysIleValIleThrGlyIleValAspAsnVal 200  
QY 535 CGCAGCTGGCTGGAGCATGCTGCTGCGGGACACCGGAGGGACTCATGCCAGGC 594  
Db 201 HisIleValArgAspAspIleLeuCysIlaGlyAsnGluLahIleAspSerCysGly 220  
QY 595 GACTCGGGAGGCCCTGGTGGCAAGGGAATGCGCACCTGCTGCAAGGGGGGGTC 654  
Db 221 AspSerGlyGlyProLeuValCysIysValGluAspThrProIleGlnAlaGlyIva 240  
QY 655 AGCTGGGGCAGGGGCTGTCGCCAACGGGCTGGATCTCACCGGTGTCACCTAC 714  
Db 241 SerTrpGlyGlyGlyCysAlaGlnProAsnArgProGlyIleValThrArgValThr 260  
QY 715 TACTTGGACTGATGCCACCATATCCCCAA 747



DE DT 16-OCT-2001 (REL 40, Last annotation update) (RMCP-7) (Tryptase, Mast cell protease 7 precursor (EC 3.4.21.59))  
 DE skin).  
 DE MCPT7 OR MCP7.  
 DE Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OC NCB: "TAXID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley;  
 RC MEDLINE=97149430; PubMed=8996238;  
 RA Lutzelschwab C., Pejler G., Aveskogh M., Hellman L.;  
 RT "secretory granule proteases in rat mast cells. Cloning of 10  
 RT different serine proteases and a carboxypeptidase A from various rat  
 RT mast cell populations";  
 J. Exp. Med. 185:13-29(1997).  
 RN [2]  
 RP SEQUENCE OF 29-53.  
 RC STRAIN=Sprague-Dawley; TISSUE=Skin;  
 RC MEDLINE=91244200; PubMed=2036367;  
 RA Braganza V.J., Simmons W.H.;  
 RT "tryptase from rat skin: purification and properties.>";  
 RL Biochemistry 30:4997-5007(1991).  
 RN [3]  
 RP SEQUENCE OF 29-51.  
 RC TISSUE=Breast carcinoma;  
 RC MEDLINE=9221826; PubMed=1314562;  
 RA ETO I., Grubbs C.J.;  
 RT "separation, purification and N-terminal sequence analysis of a novel  
 leupeptin sensitive serine endopeptidase present in chemically  
 induced rat mammary tumour.>";  
 RL Biochem. J. 281:209-216(1992).  
 CC -!- FUNCTION: TRYPTASE IS THE MAJOR NEUTRAL PROTEASE PRESENT IN MAST  
 CELLS, AND IS SECRETED UPON THE COUPLED ACTIVATION-DEGRANULATION  
 CC RESPONSE OF THIS CELL TYPE.  
 CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-!, Lys-!, but  
 CC with more restricted specificity than trypsin.  
 CC -!- SUBUNIT: HOMOTRIMER.  
 CC -!- SUBCELLULAR LOCATION: RELEASED FROM THE SECRETORY GRANULES UPON  
 CC MAST CELL ACTIVATION.  
 CC -!- TISSUE SPECIFICITY: MAST CELLS.  
 CC -!- PIM: GLYCOSYLATED (PROBABLE).  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY SI. TRYPTASE SUBFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 the European Bioinformatics Institute. There are no restrictions on  
 use by non-profit institutions, as long as its content is in no  
 modified and this statement is not removed. Usage by and for commer-  
 cial entities requires a license agreement (see <http://www.isb-sib.ch/announ->  
 or send an email to license@isb-sib.ch).

---

DR EMBL; U67910; ARB48263\_1; -  
 DR PIR; A23698; A23698.  
 DR PIR; S21275; S21275.  
 DR HSSP; P20231; 1AAO.  
 DR MEROPS; S01\_026; -.  
 DR MEROPS; S01\_143; -.  
 DR InterPro; IPRO01314; Chymotrypsin.  
 DR InterPro; IPRO01254; Ser\_protease\_Try.  
 DR PRIM; PF00059; trypsin; 1.  
 DR PRINTS; PRO0722; CHYMOTRYPSIN.  
 DR SMART; SM00220; TRYPSIN\_SPC; 1.  
 DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SBR; 1.  
 KW Hydrolase; Serine protease; Glycoprotein; Zymogen; Signal;  
 KW Multigene family.  
 FT SIGNAL; 1 18 POTENTIAL.  
 FT PROPEP 19 28 ACTIVATION PEPTIDE.  
 FT CHAIN 29 273 MAST CELL PROTEASE 7.  
 FT ACT\_SITE 72 72 MAST CELL PROTEASE SYSTEM (BY SIMILARITY).

FT	ACT_SITE	119	119	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE	222	222	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	DISULFID	57	73	BY SIMILARITY.
FT	DISULFID	153	228	BY SIMILARITY.
FT	DISULFID	186	209	BY SIMILARITY.
FT	DISULFID	218	246	BY SIMILARITY.
FT	CARBHYD	49	49	N-LINKED (GLICNAc. . .) (PROBABLE).
FT	CONFLICT	42	42	W > Y (IN REF. 3).
FT	CONFLICT	49	51	NDT -> WLP (IN REF. 3).
SQ	SEQUENCE	273 AA:	30400 MW:	65A5ED4D279FB284 CRC64;
Alignment Scores:				
Prod. No.:		9.98e-54	Length:	273
Score:		1039.00	Matches:	188
Percent Similarity:		84.27%	Conservative:	21
Best Local Similarity:		75.81%	Mismatches:	38
Query Match:		71.26%	Indels:	1
DB:		1	Gaps:	0
US-09-598-982-20 (1-771) x MCT7-RAT (1-273)				

RESULT 10  
**TRYT\_SHEEP** STANDARD; PRT; 273 AA.  
ID TRYT\_SHEEP  
AC Q9XSN2;  
DT 16-OCT-2001 (Rel. 40, created)  
DT 16-OCT-2001 (Rel. 40, last sequence update)  
DT 15-JUN-2002 (Rel. 41, last annotation update)  
DB Tryptase 2 precursor (EC 3.4.21.5).  
Ox Ovis aries (Sheep).  
Oc Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Fecora; Bovidae; Bovidae; Caprinae; Ovis.  
OC NCBI\_TaxID=9940;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Abomasum;  
RX MEDLINE=20308142; PubMed=10848900;  
RA Pemberton A.D., McLeese S.M., Huntley J.F., Collie D.D.S., Scudamore C.L., McGuire A.R., Walls A.F., Miller H.R.P.;  
RT "cDNA sequence of two sheep mast cell tryptases and the differential expression of trypase and sheep mast cell proteinase-1 in lung, dermis and gastrointestinal tract";  
RL Clin. Exp. Allergy 30:818-832(2000).  
CC -- FUNCTION: TRYPTASE IS THE MAJOR NEUTRAL PROTEASE PRESENT IN MAST CELLS AND IS SECRETED UPON THE COUPLED ACTIVATION-DEGRANULATION RESPONSE OF THIS CELL TYPE.  
CC -- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-, Lys-|-, but with more restricted specificity than trypsin.  
CC -- SUBUNIT: HOMOTRIMER (BY SIMILARITY).  
CC -- CELLULAR LOCATION: RELEASED FROM THE SECRETORY GRANULES UPON MAST CELL ACTIVATION.  
CC -- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. TRYPTASE SUBFAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).  
CC  
DR EMBL: Y18224; CAB41989.1; -.  
DR HSSP: P20331; IAAO.  
DR MEROPS: S01\_143; -.  
DR InterPro: IPR001314; Chymotrypsin.  
DR InterPro: IPR001254; Ser\_protease\_Try.  
DR Pfam: PF00089; trypsin; 1.  
DR PRINTS: PR0072; CHYMOPTRPSIN.  
DR SMART: SM00020; TRYPT\_SPC; 1.  
DR PROSITE: PS50240; TRYPT\_DOM; 1.  
DR PROSITE: PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE: PS00135; TRYPSIN\_SER; 1.  
DR Hydrolase; Serine protease; Signal; Glycoprotein; Zymogen.  
FT SIGNAL 1 18 POTENTIAL.  
FT PROPEP 19 28 ACTIVATION\_PEPIDE (BY SIMILARITY).  
FT CHAIN 29 273 TRYPTASE 2.  
FT ACT\_SITE 72 72 CHARGE\_RELAY\_SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 119 119 CHARGE\_RELAY\_SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 222 222 CHARGE\_RELAY\_SYSTEM (BY SIMILARITY).  
FT DISULFID 57 73 BY SIMILARITY.  
FT DISULFID 153 228 BY SIMILARITY.  
FT DISULFID 186 209 BY SIMILARITY.  
FT DISULFID 218 246 BY SIMILARITY.  
FT CARBOHYD 231 231 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ 273 AA; 30288 MW; DEBA79218C3B67D CFC64;

Alignment Scores:  
Pred. No.: 3.33e-53 Length: 273  
Score: 1030.00 Matches: 179  
Percent Similarity: 84.08% Conservative: 27  
Best Local Similarity: 73.05% Mismatches: 39  
Query Match: 70.648 Indels: 0  
Gaps: 0

US-09-598-982-20 (1-771) x TRYT\_SHEEP (1-273)  
RESULT 11  
**TRYT\_PIG** STANDARD; PRT; 275 AA.  
ID TRYT\_PIG  
AC Q9N2D1;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, last sequence update)  
DT 15-JUN-2002 (Rel. 41, last annotation update)  
DB Tryptase precursor (EC 3.4.21.59).  
GN MCT7.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OC NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lung;  
RX MEDLINE=20285343; PubMed=10824103;  
RA Chen Y., Shiota M., Ohuchi M., Towattari T., Tashiro J., Murakami M.,





FT CARBOHYD 95 95 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 106 106 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 269 AA; 29824 MW; B4A4AF7E4E2FD4E5 CRC64;

**Alignment Scores:**

Pred. No.:	4.29e-33	Length:	269
Score:	684.50	Matches:	132
Percent Similarity:	65.34%	Conservative:	32
Best Local Similarity:	52.59%	Mismatches:	78
Query Match:	46.95%	Indices:	9
DB:	1	Gaps:	4

US-09-598-982-20 (1-771) x TRYM\_CANFA (1-269)

QY 19 ATCCTCGGGGGTCAGGAGCCCCAGGACCAGTGCCCTGGCAGGTGAGAGTC 78  
 |||||||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:  
 Db 20 IleValGlyGlyCysLysValProAlaArgArgTyrProTrpGlnValSerLeuArgPhe 39

QY 79 CACGCC-----CCATRACTSGTACTCTGCGGGGTCCTCATCCACCCCCAG 129  
 |||||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:  
 Db 40 HisGlyMetGlySerGlyGlnTrpGlnHisLeuCysGlySerLeuLeuHisPrgIn 59

QY 130 TGGTGTCTGACCGCCGGCGCTGGGGACCGAACGATCTGGCCGCTCAGG 189  
 |||||||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:  
 Db 60 TrpValLeuThrAlaAlaHisCysValGluLeuGlyLeuGluAlaAlaThrLeuArg 79

QY 190 GTGCAACTGGGGAGCAGGACCTCTACTACCAAGGACCACTGCTGCCGTCAGAGGATC 249  
 |||||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:  
 Db 80 ValGlnValGlyGlnLeuArgLeuArgLeuTyrAspHisAspGlnLeuCysAsnValThrGluIle 99

QY 250 ATCCTGCACCAACGATTCTCACACGCCAGATCGA-----GCGACATGCCG 300  
 |||||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:  
 Db 100 IleGlyIleProAsnPheAsnMetSerTrpTyrGlyTyrPaspThrAlaAspIleAlaLeu 119

QY 301 CTGAGCTGGAGGAGCCGGTGAAGGCTCCAGCACGTCACCGTCACTGCCCT 360  
 |||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:  
 Db 120 LeuIleLeuGluIlaProLeuThrIleUserGluAspValAsnLeuValSerLeuProSer 139

QY 361 GCTCTAGAGACTCCCGGGGATGCCGCTGCTGGTACTGGCTGGGGATGTTGGAC 420  
 |||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:  
 Db 140 ProSerLeuIleValProGlyMetLeuCysTrpValThrGlyTrpIleAspIleAla 159

QY 421 AATGAGGAGGCCCTCCACGCCGATTCCTGAGGAGGTGCCATATGGAA 480  
 |||||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:  
 Db 160 AspAspIleProLeuProIleProTyrIleLeuGluIleValProIleValGly 179

QY 481 AACCCATCTGGAGCCAAATAACCCACCTGGCGCTACACGGGACACGGCAGTC 540  
 |||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:  
 Db 180 AsnArgGlyCysAsnCysHistYtGln----ThrIleLeuGluGlnAspBgluVal 197

QY 541 GTCCTGGACACACTGCTGCTGGGAAACACCGGAGGACTCATGCCGGGACTCC 600  
 |||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:  
 Db 198 IleLysginIleAspMetLeuCysAlaGlySerGlyIleHisAspSerCysGlnMetAspSer 217

QY 601 GGAGGCCCTGGAGCAGGAGGACCTGGGACGGGGGGGGTGTGAGCTGG 660  
 |||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:  
 Db 218 GlyGlyProLeuValCysArgTrpLysCysThrTrpIleGlnValGlyValSerTrp 237

QY 661 GGCAGGGGCTGGCCAGGCCAACGGGCTGGCATCACCCGTGTCACTACTG 720  
 |||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:  
 Db 238 GlyTyrGlyGlyGlyTyr--AsnLeuProGlyIleValAlaArgValThrSerIleVal 256

QY 721 GACTGATCCACCACTATGTCCTCAACCCGTGTCACTACTG 753  
 |||:::|||:::|||:::|||:::|||:::|||:::|||:  
 Db 257 SerTrpIleHisGlnIleProLeuSerPro 267

RESULT 14

TRY-MOUSE STANDARD; PRT; 311 AA.

ID TRY-MOUSE STANDARD; PRT; 311 AA.

AC 090017; 16-AAGATCGTCGGGGTCAGGGCCCCAGGACAGTGCCCTGGCAGGTGAGCTGAGA 75  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-Oct-2001 (Rel. 40, Last annotation update)

DE TRYPTASE gamma precursor (EC 3.4.21.-) (Transmembrane tryptase).  
 GN TP5G1 OR TM.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI-TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/Sv, and BALB/C;  
 RR MEDLINE=9452074; PubMed=10521469;  
 RA Wong G.W., Tang Y., Feyrant E., Sali A., Li L., Li Y., Huang C., Friend D.S., Krilis S.A., Stevens R.L.;  
 RA "Identification of a new member of the tryptase family of mouse and human mast cell proteases which possesses a novel COOH-terminal hydrophobic extension.";  
 CC J. Biol. Chem. 274:30784-30793 (1999).  
 CC -!- SUBCELLULAR LOCATION: MEMBRANE-ANCHORED (Potential).  
 CC -!- TISSUE SPECIFICITY: Expressed in many tissues.  
 CC -!- SIMILARITY: BELONGS TO PEPTIDE FAMILY SI. TRYPTASE SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; AF175760; AAO03698\_1; -.  
 DR EMBL; AF17523; AAO03696\_1; -.  
 DR HSSP; P2031; IAO.  
 DR MEGP; S01\_028; -.  
 DR MGDB; MGI:134991; TP5G1.  
 DR InterPro; IPR0011314; Chymotrypsin.  
 DR Pfam; PF00089; TRYPSIN.  
 DR SMART; SM0020; TRYPSIN; 1.  
 DR PROSITE; PS50240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_HIS; FALSE\_NBG.  
 KW Hydrolase; Serine protease; Signal; Glycoprotein; Zymogen; Transmembrane.  
 KW SIGNAL 1 16 POTENTIAL.  
 FT CHAIN 17 28 TRYPTASE GAMMA LIGHT CHAIN.  
 FT CHAIN 30 311 TRYPTASE GAMMA HEAVY CHAIN.  
 FT TRANSEM 277 297 POTENTIAL.  
 FT ACT\_SITE 70 70 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 117 117 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 214 214 INTERCHAIN (POTENTIAL).  
 FT DISULFID 18 137 BY SIMILARITY.  
 FT DISULFID 55 71 BY SIMILARITY.  
 FT DISULFID 151 220 BY SIMILARITY.  
 FT DISULFID 184 202 BY SIMILARITY.  
 FT DISULFID 210 238 BY SIMILARITY.  
 FT CARBOHYD 77 77 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 192 192 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 311 AA; 32656 MW; 7FC9D6EF6A2AB808 CRC64;

Alignment scores:

Pred. No.:	1.13e-28	Length:	311
Score:	608.50	Matches:	123
Percent Similarity:	60.63%	Conservative:	31
Best Local Similarity:	48.43%	Mismatches:	81
Query Match:	41.74%	Indices:	19
DB:	1	Gaps:	6

US-09-598-982-20 (1-771) x TRY-MOUSE (1-311)

QY 16 AGAATCGTCGGGGTCAGGGCCCCAGGACAGTGCCCTGGCAGGTGAGCTGAGA 75  
 |||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:  
 Db 29 ArgIleValGlyGlyIleAlaProAlaGlyIleValProTrpProGlnIleLeuArg 48

Qy 76 GTCGCCAGCCATACATGGATGCACTTCGGGGCCTCCATCCCCAGTGGTG 135  
   :::||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 49 LeuThrAlaHisCysPheSerValAsnSer--SerAspTygInWalHs 84  
   |||||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Qy 136 CTGACCGCGCCGGGGTGTGGACGGACGTCAAGGACTGGCCCTCAGGGTCAA 195  
   |||||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Db 66 LeuThrAlaHisCysPheSerValAsnSer--SerAspTygInWalHs 84  
   |||||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Qy 196 CTGCGGGACAGCACCTRACTACAGGACCAGCTGCAGTCAGCAGGATCATCGG 255  
   |||||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Db 85 LeuGlyGluLeuThrValThrLeuSerProHisPheSerThrValLysArgIleMet 104  
   |||||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Qy 256 CACCCACAGTCTACACGCCAGATCGA-----GGGACATCCCCCTG 300  
   |||||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Db 105 -----TyThrGlySerProGlyProProGlySerSerGlyAspIleAlaLeu 120  
   |||||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Qy 301 CTGGAGCTGGAGGAGCCGGTGAAGGCTCTCAGGTCACCGTACCCCTGCCCT 360  
   |||||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Db 121 ValGlnLeuSerSerProValAlaLeuSerSerInValInProValysLeuProIu 140  
   |||||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Qy 361 GGCAGTAGAGACCTTCCCAGGGGAGCCGCTGCGTGGTCACTGSGCTGCGATGGAC 420  
   |||||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Db 141 AlasSerAlaAspPheTerProGlyMetGlnCystPheTerGlyTrpGlyTerGly 160  
   |||||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Qy 421 AATGATGACCGCCCTCCACGCCATTCTCTGAAGCAGGTGAAGGCTCCATAATGGAA 480  
   |||||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Db 161 GluGlyGluProLeuLeuProProTerProTerAsnLeuGluLalaLysValSerValAsp 180  
   |||||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Qy 481 AACCCATTTGAGCCAAATAACCAACCTGGGCTAACGGGAGCAGCTGGCGCT 540  
   |||||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Db 181 ValYsThrCysSerGln-----AlaTyrsAsnSerProAsnGlySerLeu 195  
   |||||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Qy 541 GTCCCCGACACATCTGTTGCGGGACACCCGGAGGACTCATGCCAGGCGACTCC 600  
   |||||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Db 196 IleGlnProAspMetLeuCysAlaArgGlyPro--GlyAspAlaCysGlnAspAspSer 214  
   |||||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Qy 601 GGAGGGCCCTGGTGTGCAAGGTCATGCCACCGCGCTGCAGGGGGCGGGCAGCTGG 660  
   |||||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Db 215 GlyLyGlyProLeuValCysGlnValAlaGlyThrPrngInAlaGlyValSerTrp 234  
   |||||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Qy 661 GGCAGGGCTGTGCCAGCCAAACGGGCTGCCATCTAACCCGGTCACTACTTG 720  
   |||||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Db 235 GlyGluGlyCysGlyArgProAspArgProGlyValTerAlaArgValThrAlaTerVal 254  
   |||||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 RESULT 15  
 TRY\_G\_HUMAN STANDARD; PRT; 321 AA.  
 ID TRY\_G\_HUMAN STANDARD; PRT; 321 AA.  
 AC 09NR82; 09NR08; 096015; 09UBB2;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE 16-OCT-2001 (Rel. 40, Last annotation update)  
 TP5G1 OR TMT.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_TAXID=9606;  
 RN [1] SEQUENCE FROM N.A. (VARIANTS GAMMA-1 AND GAMMA-2).  
 RX MEDLINE=2013-02-28;3; PubMed=10843716;  
 RA Caughey G.H., Raymond W.W., Blount J.L., Hau L.W., Pallaoro M.,  
 RA Wolters P.J., Vergheze G.M.;  
 RT "Characterization of human gamma-trypases, novel members of the  
 chromosome 16p mast cell tryptase and prostatin gene families.",  
 J. Immunol. 164:6566-6575(2000).  
 RX SEQUENCE FROM N.A.  
 RA MEDLINE=9452974; PubMed=10521469;  
 RX Wong G.W., Tang Y., Feyfant E., Sali A., Li L., Li Y., Huang C.,

Ra Friend D.S., Krilis S.A., Stevens R.L.;  
 Rr "Identification of a new member of the trypase family of mouse and  
 human mast cell proteases which possesses a novel COOH-terminal  
 hydrophobic extension.";  
 Rr J. Biol. Chem. 274:30784-30793(1999).  
 Rr [3]  
 RP SEQUENCE OF 220-321 FROM N.A.  
 RA Mittman S., Agnew W.S.;  
 RR "Organization and alternative splicing of CACNA1H";  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.  
 CC -!- SUBCELLULAR LOCATION: MEMBRANE-ANCHORED (Potential).  
 CC -!- TISSUE-SPECIFICITY: Expressed in many tissues  
 CC differ by 5 residues.  
 CC -!- POLYMORPHISM: There are two alleles; gamma-I and gamma-II which  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@sb-sib.ch).  
 CC  
 DR EMBL; AF91031; AAP76457.1; -  
 DR EMBL; AF195508; AAFT6458.1; -  
 DR EMBL; AF175759; AAFO3697.1; -  
 DR EMBL; AF17552; AAFO3695.1; -  
 DR EMBL; AF223563; AAGA8852.2; -  
 DR HSSP; P00763; IEDB.  
 DR MEROPS; S01\_028; -  
 DR Genew; HGNC;14134; TPSGL.  
 DR InterPro; IPR001314; Chymotrypsin.  
 DR InterPro; IPR00125; Ser\_Protease\_Try.  
 DR PRAM; PF00089; trypsin; 1.  
 DR PRAM; P00722; CHYMOTRYPSIN.  
 DR SMART; SM00020; TRY\_P\_SCI; 1.  
 DR PROSITE; PS50240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 KW Hydrolase; Serine protease; Signal; Glycoprotein; Zymogen;  
 KW Transmembrane; Polymorphism.  
 FT SIGNAL 1 19  
 FT CHAIN 20 36  
 FT CHAIN 38 321  
 FT TRANSEM 284 304  
 FT ACT\_SITE 78 78  
 FT ACT\_SITE 125 125  
 FT ACT\_SITE 222 222  
 FT DISUFRID 26 145  
 FT DISUFRID 63 79  
 FT DISUFRID 159 228  
 FT DISUFRID 192 210  
 FT DISUFRID 218 246  
 FT CARBOHYD 85 85  
 FT VARIANT 60 60  
 FT VARIANT 126 126  
 FT VARIANT 132 132  
 FT VARIANT 132 132  
 FT VARIANT 204 204  
 FT VARIANT 288 288  
 FT VARIANT 160 160  
 SQ SEQUENCE 321 AA; 33827 MW; FFP7B06E3C4A962D CRC64;  
 Alignment Scores:  
 Pred. No.: 8.39e-28 Length: 321  
 Score: 593.50 Matches: 119  
 Percent Similarity: 61.89% Conservative: 32  
 Best Local Similarity: 48.77% Mismatches: 80

Query Match:	40.71%	Indels:	13
DB:	1	Gaps:	6
Qy	16 AGATCGGGGGTCAGGAGGCCAGGAGAATGAGCCCTGGCAAGTGACCTGTAGA		75
Db	37 ArgilevalGlyGlyHISAlaAlaProAlaGlyAlanirProTrpGinAlaserLeuarg		56
Qy	76 GTCCACGCCATACTGCATGCACTCTGGGGGCGGCTCCTCACCCCCAGGGTG		135
Db	57 LeuArgArg-----MethIleValCysGlyLyserLeuLeuUser-ProGlnTrpIal		73
Qy	136 CTGACGCCGGGGCTGGGACGGACTGCAAGGATCTGGCCCTCAGGTGCAA		195
Db	74 LeuthAlaAlaHisCysPheserGlyserLeuAsnSer--SerAspTyrrGlnValHis		92
Qy	196 CTCGGGAGCAGCAGCTCTACTACCCAGGACCAGCTGCTGGGAGCAGAGGATCATCCTG		255
Db	93 LeuGlyGluIleGluIleThrLeuSerProHisPheSerThrValArgGlnIleIleIeu		112
Qy	256 CACCCACAGTCTACACCCCCAGATCGA-----GGGACATGCCCTGCGAGCTG		309
Db	113 HisSerSer---ProSerGlyIleThrSerGlyAspIleAlaLeuValGluIeu		131
Qy	310 GAGGGCCGCTGAAGTCRCCAGGCCACAGGTACCCCTGCCCTGCCTCAGAG		369
Db	132 SerValProValThrIleUserSerArgIleLeuProValCysLeuProGluIalAsp		151
Qy	370 ACCCTCCCCGGGGATGCCGTGCTGGGTAACCTGGCTGGGGGAGATGATGAG		429
Db	152 AspPheCysProGlyIleArgCysTrpValThrGlyTrpGlyYrrThrArgGluGlyGu		171
Qy	430 CGCCCTCCACCGCCATTCTCTGAGCAGGTGGGCCATGCAAACCCACAMP		489
Db	172 ProLeuProProProTyrsSerLeuArgGluValIysValSerValVaAspThrGluThr		191
Qy	490 TGTGAGCGAAATACACCTTGGCCCTACACGGAGAGACGCCATGTTGGTAC		549
Db	192 CysArgArgAsp-----TyrProGlyProGlySerIleLeuGlnPro		206
Qy	550 GACATGCTGTCGGGGAACACCGGGAGGACTCATGCCAGGGACTCCGGAGGCC		609
Db	207 AspMetLeuCysAlaArgGlyPro---GlyAspAlaCysGlnAspAspSerGlyGlyPro		225
Qy	610 CTGGCTGCGAGGGTCAAGGCTGGCTGCAGGGGGCTGGCTGGCTGGGGGGAGGC		669
Db	226 LeuValCysGlnValAsnGlyAlaTrpValGlnIleValSerTrpGlyGluGly		245
Qy	670 TGTGCCACGCCAACGGCTGGCATCTACACCGGTGACCTACTACTTGACTGGATC		729
Db	246 CysGlyArgProAsnArgProGlyValTyrThrArgValProAlaTyrValAsnTrpIle		265
Qy	730 CACCACTATGTC 741		
Db	266 ArgArgHisIle 269		

Search completed: January 31, 2003, 06:54:08  
 Job time : 34.5 secs

Gencore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model  
Run on: January 31, 2003, 06:53:20 ; Search time 16.5 Seconds  
(without alignments)  
2749.706 Million cell updates/sec

Title: US-09-598-982-20  
Perfect score: 1458

Sequence: 1. 999ccctcgagaaaaaat.....cgtgaaggccgcgtcgt 771

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 262574 seqs, 2942292 residues

Total number of hits satisfying chosen parameters: 525148  
Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

```
-MODEL=frame+ n2p model -DEY=x1h
-Q=/cgn2-1/USPRO_spool/US095982/runat_27012003_073718_19604/app_query.fasta_1.967
-DB=Issued_Patents_AA -QFMT=fastaN -SUFFIX=n2p.rai -MNMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -STAR=1 -END=-1 -MATRIX=Blosum62 -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR_SCORE=PCT -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=PTO -NORM=ext -HEARSPINE=5000 -MINLEN=0 -MAXLEN=100000000
-USR=US095982_ECGN_1_4 @runat_27012003_073718_19604 -NCPU=6 -ICPU=3
-NO_XLPIX -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database : Issued\_Patents\_AA;\*

```
1: /cgn2_6/podata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/podata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/podata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/podata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/podata/1/iaa/PCUS_COMB.pep:*
6: /cgn2_6/podata/1/iaa/backfiles1.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

**RESULT 1**  
US-09-598-970A-5  
Sequence 5, Application US/09079970A  
; Sequence 5, Application US/09079970A  
; Patent No. 6274366  
; GENERAL INFORMATION:  
; APPLICANT: Maffitt, Mark A.  
; APPLICANT: Niles, Andrew L.  
; APPLICANT: Haak-Freischl, Mary  
; TITLE OF INVENTION: Enzymatically-Active Recombinant Human  
; TITLE OF INVENTION: Beta-Tryptase and Method of Making Same  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Intellectual Property Department  
; STREET: 8000 Exceislor Drive, Suite 401  
; CITY: Madison  
; STATE: WISCONSIN  
; COUNTRY: U.S.A.  
; ZIP: 53717-1914  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY disk

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/079, 970A  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatternIn Release #1.0, Version #1.30  
APPLICATION NUMBER: US/09/079, 970A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Leone, Joseph T.  
REGISTRATION NUMBER: 37, 170  
REFERENCE/DOCKET NUMBER: 34506 . 073  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (608) 831-2100  
TELEFAX: (608) 831-2106  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:

Result No.	Score	% Match	Query Length	DB ID	Description
1	1387	95.1	249	4	US-09-079-970A-5 Sequence 5, Appli Sequence 6, Appli Sequence 21, Appli Sequence 16, Appli Sequence 19, Appli Sequence 14, Appli Sequence 23, Appli Sequence 18, Appli Sequence 69, Appli Sequence 17, Appli Sequence 12, Appli Sequence 15, Appli
2	1368	93.8	245	4	US-09-079-970A-6 Sequence 1, Appli Sequence 2, Appli Sequence 3, Appli Sequence 4, Appli Sequence 5, Appli Sequence 6, Appli Sequence 7, Appli Sequence 8, Appli Sequence 9, Appli Sequence 10, Appli Sequence 11, Appli Sequence 12, Appli Sequence 13, Appli Sequence 14, Appli Sequence 15, Appli Sequence 16, Appli Sequence 17, Appli Sequence 18, Appli Sequence 19, Appli Sequence 20, Appli Sequence 21, Appli Sequence 22, Appli Sequence 23, Appli Sequence 24, Appli Sequence 25, Appli Sequence 26, Appli Sequence 27, Appli Sequence 28, Appli Sequence 29, Appli Sequence 30, Appli Sequence 31, Appli Sequence 32, Appli Sequence 33, Appli Sequence 34, Appli Sequence 35, Appli Sequence 36, Appli Sequence 37, Appli Sequence 38, Appli Sequence 39, Appli Sequence 40, Appli Sequence 41, Appli Sequence 42, Appli Sequence 43, Appli Sequence 44, Appli Sequence 45, Appli
3	1368	93.8	274	2	US-09-079-970A-21 Sequence 1, Appli Sequence 2, Appli Sequence 3, Appli Sequence 4, Appli Sequence 5, Appli Sequence 6, Appli Sequence 7, Appli Sequence 8, Appli Sequence 9, Appli Sequence 10, Appli Sequence 11, Appli Sequence 12, Appli Sequence 13, Appli Sequence 14, Appli Sequence 15, Appli Sequence 16, Appli Sequence 17, Appli Sequence 18, Appli Sequence 19, Appli Sequence 20, Appli Sequence 21, Appli Sequence 22, Appli Sequence 23, Appli Sequence 24, Appli Sequence 25, Appli Sequence 26, Appli Sequence 27, Appli Sequence 28, Appli Sequence 29, Appli Sequence 30, Appli Sequence 31, Appli Sequence 32, Appli Sequence 33, Appli Sequence 34, Appli Sequence 35, Appli Sequence 36, Appli Sequence 37, Appli Sequence 38, Appli Sequence 39, Appli Sequence 40, Appli Sequence 41, Appli Sequence 42, Appli Sequence 43, Appli Sequence 44, Appli Sequence 45, Appli
4	1368	93.8	274	2	US-09-079-970A-21 Sequence 1, Appli Sequence 2, Appli Sequence 3, Appli Sequence 4, Appli Sequence 5, Appli Sequence 6, Appli Sequence 7, Appli Sequence 8, Appli Sequence 9, Appli Sequence 10, Appli Sequence 11, Appli Sequence 12, Appli Sequence 13, Appli Sequence 14, Appli Sequence 15, Appli Sequence 16, Appli Sequence 17, Appli Sequence 18, Appli Sequence 19, Appli Sequence 20, Appli Sequence 21, Appli Sequence 22, Appli Sequence 23, Appli Sequence 24, Appli Sequence 25, Appli Sequence 26, Appli Sequence 27, Appli Sequence 28, Appli Sequence 29, Appli Sequence 30, Appli Sequence 31, Appli Sequence 32, Appli Sequence 33, Appli Sequence 34, Appli Sequence 35, Appli Sequence 36, Appli Sequence 37, Appli Sequence 38, Appli Sequence 39, Appli Sequence 40, Appli Sequence 41, Appli Sequence 42, Appli Sequence 43, Appli Sequence 44, Appli Sequence 45, Appli
5	1363	93.5	273	2	US-09-079-970A-19 Sequence 1, Appli Sequence 2, Appli Sequence 3, Appli Sequence 4, Appli Sequence 5, Appli Sequence 6, Appli Sequence 7, Appli Sequence 8, Appli Sequence 9, Appli Sequence 10, Appli Sequence 11, Appli Sequence 12, Appli Sequence 13, Appli Sequence 14, Appli Sequence 15, Appli Sequence 16, Appli Sequence 17, Appli Sequence 18, Appli Sequence 19, Appli Sequence 20, Appli Sequence 21, Appli Sequence 22, Appli Sequence 23, Appli Sequence 24, Appli Sequence 25, Appli Sequence 26, Appli Sequence 27, Appli Sequence 28, Appli Sequence 29, Appli Sequence 30, Appli Sequence 31, Appli Sequence 32, Appli Sequence 33, Appli Sequence 34, Appli Sequence 35, Appli Sequence 36, Appli Sequence 37, Appli Sequence 38, Appli Sequence 39, Appli Sequence 40, Appli Sequence 41, Appli Sequence 42, Appli Sequence 43, Appli Sequence 44, Appli Sequence 45, Appli
6	1363	93.5	273	2	US-09-079-970A-14 Sequence 1, Appli Sequence 2, Appli Sequence 3, Appli Sequence 4, Appli Sequence 5, Appli Sequence 6, Appli Sequence 7, Appli Sequence 8, Appli Sequence 9, Appli Sequence 10, Appli Sequence 11, Appli Sequence 12, Appli Sequence 13, Appli Sequence 14, Appli Sequence 15, Appli Sequence 16, Appli Sequence 17, Appli Sequence 18, Appli Sequence 19, Appli Sequence 20, Appli Sequence 21, Appli Sequence 22, Appli Sequence 23, Appli Sequence 24, Appli Sequence 25, Appli Sequence 26, Appli Sequence 27, Appli Sequence 28, Appli Sequence 29, Appli Sequence 30, Appli Sequence 31, Appli Sequence 32, Appli Sequence 33, Appli Sequence 34, Appli Sequence 35, Appli Sequence 36, Appli Sequence 37, Appli Sequence 38, Appli Sequence 39, Appli Sequence 40, Appli Sequence 41, Appli Sequence 42, Appli Sequence 43, Appli Sequence 44, Appli Sequence 45, Appli
7	1344	92.2	267	2	US-09-079-970A-23 Sequence 1, Appli Sequence 2, Appli Sequence 3, Appli Sequence 4, Appli Sequence 5, Appli Sequence 6, Appli Sequence 7, Appli Sequence 8, Appli Sequence 9, Appli Sequence 10, Appli Sequence 11, Appli Sequence 12, Appli Sequence 13, Appli Sequence 14, Appli Sequence 15, Appli Sequence 16, Appli Sequence 17, Appli Sequence 18, Appli Sequence 19, Appli Sequence 20, Appli Sequence 21, Appli Sequence 22, Appli Sequence 23, Appli Sequence 24, Appli Sequence 25, Appli Sequence 26, Appli Sequence 27, Appli Sequence 28, Appli Sequence 29, Appli Sequence 30, Appli Sequence 31, Appli Sequence 32, Appli Sequence 33, Appli Sequence 34, Appli Sequence 35, Appli Sequence 36, Appli Sequence 37, Appli Sequence 38, Appli Sequence 39, Appli Sequence 40, Appli Sequence 41, Appli Sequence 42, Appli Sequence 43, Appli Sequence 44, Appli Sequence 45, Appli
8	1344	92.2	267	2	US-09-079-970A-23 Sequence 1, Appli Sequence 2, Appli Sequence 3, Appli Sequence 4, Appli Sequence 5, Appli Sequence 6, Appli Sequence 7, Appli Sequence 8, Appli Sequence 9, Appli Sequence 10, Appli Sequence 11, Appli Sequence 12, Appli Sequence 13, Appli Sequence 14, Appli Sequence 15, Appli Sequence 16, Appli Sequence 17, Appli Sequence 18, Appli Sequence 19, Appli Sequence 20, Appli Sequence 21, Appli Sequence 22, Appli Sequence 23, Appli Sequence 24, Appli Sequence 25, Appli Sequence 26, Appli Sequence 27, Appli Sequence 28, Appli Sequence 29, Appli Sequence 30, Appli Sequence 31, Appli Sequence 32, Appli Sequence 33, Appli Sequence 34, Appli Sequence 35, Appli Sequence 36, Appli Sequence 37, Appli Sequence 38, Appli Sequence 39, Appli Sequence 40, Appli Sequence 41, Appli Sequence 42, Appli Sequence 43, Appli Sequence 44, Appli Sequence 45, Appli
9	1274	87.4	245	4	US-09-079-944-483-69 Sequence 1, Appli Sequence 2, Appli Sequence 3, Appli Sequence 4, Appli Sequence 5, Appli Sequence 6, Appli Sequence 7, Appli Sequence 8, Appli Sequence 9, Appli Sequence 10, Appli Sequence 11, Appli Sequence 12, Appli Sequence 13, Appli Sequence 14, Appli Sequence 15, Appli Sequence 16, Appli Sequence 17, Appli Sequence 18, Appli Sequence 19, Appli Sequence 20, Appli Sequence 21, Appli Sequence 22, Appli Sequence 23, Appli Sequence 24, Appli Sequence 25, Appli Sequence 26, Appli Sequence 27, Appli Sequence 28, Appli Sequence 29, Appli Sequence 30, Appli Sequence 31, Appli Sequence 32, Appli Sequence 33, Appli Sequence 34, Appli Sequence 35, Appli Sequence 36, Appli Sequence 37, Appli Sequence 38, Appli Sequence 39, Appli Sequence 40, Appli Sequence 41, Appli Sequence 42, Appli Sequence 43, Appli Sequence 44, Appli Sequence 45, Appli
10	1274	87.4	275	2	US-09-079-944-483-69 Sequence 1, Appli Sequence 2, Appli Sequence 3, Appli Sequence 4, Appli Sequence 5, Appli Sequence 6, Appli Sequence 7, Appli Sequence 8, Appli Sequence 9, Appli Sequence 10, Appli Sequence 11, Appli Sequence 12, Appli Sequence 13, Appli Sequence 14, Appli Sequence 15, Appli Sequence 16, Appli Sequence 17, Appli Sequence 18, Appli Sequence 19, Appli Sequence 20, Appli Sequence 21, Appli Sequence 22, Appli Sequence 23, Appli Sequence 24, Appli Sequence 25, Appli Sequence 26, Appli Sequence 27, Appli Sequence 28, Appli Sequence 29, Appli Sequence 30, Appli Sequence 31, Appli Sequence 32, Appli Sequence 33, Appli Sequence 34, Appli Sequence 35, Appli Sequence 36, Appli Sequence 37, Appli Sequence 38, Appli Sequence 39, Appli Sequence 40, Appli Sequence 41, Appli Sequence 42, Appli Sequence 43, Appli Sequence 44, Appli Sequence 45, Appli
11	1274	87.4	275	2	US-09-079-944-483-69 Sequence 1, Appli Sequence 2, Appli Sequence 3, Appli Sequence 4, Appli Sequence 5, Appli Sequence 6, Appli Sequence 7, Appli Sequence 8, Appli Sequence 9, Appli Sequence 10, Appli Sequence 11, Appli Sequence 12, Appli Sequence 13, Appli Sequence 14, Appli Sequence 15, Appli Sequence 16, Appli Sequence 17, Appli Sequence 18, Appli Sequence 19, Appli Sequence 20, Appli Sequence 21, Appli Sequence 22, Appli Sequence 23, Appli Sequence 24, Appli Sequence 25, Appli Sequence 26, Appli Sequence 27, Appli Sequence 28, Appli Sequence 29, Appli Sequence 30, Appli Sequence 31, Appli Sequence 32, Appli Sequence 33, Appli Sequence 34, Appli Sequence 35, Appli Sequence 36, Appli Sequence 37, Appli Sequence 38, Appli Sequence 39, Appli Sequence 40, Appli Sequence 41, Appli Sequence 42, Appli Sequence 43, Appli Sequence 44, Appli Sequence 45, Appli
12	1092	74.9	276	2	US-09-079-404B-21 Sequence 1, Appli Sequence 2, Appli Sequence 3, Appli Sequence 4, Appli Sequence 5, Appli Sequence 6, Appli Sequence 7, Appli Sequence 8, Appli Sequence 9, Appli Sequence 10, Appli Sequence 11, Appli Sequence 12, Appli Sequence 13, Appli Sequence 14, Appli Sequence 15, Appli Sequence 16, Appli Sequence 17, Appli Sequence 18, Appli Sequence 19, Appli Sequence 20, Appli Sequence 21, Appli Sequence 22, Appli Sequence 23, Appli Sequence 24, Appli Sequence 25, Appli Sequence 26, Appli Sequence 27, Appli Sequence 28, Appli Sequence 29, Appli Sequence 30, Appli Sequence 31, Appli Sequence 32, Appli Sequence 33, Appli Sequence 34, Appli Sequence 35, Appli Sequence 36, Appli Sequence 37, Appli Sequence 38, Appli Sequence 39, Appli Sequence 40, Appli Sequence 41, Appli Sequence 42, Appli Sequence 43, Appli Sequence 44, Appli Sequence 45, Appli

## ALIGNMENTS

; LENGTH: 249 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-09-079-970A-5  
 Alignment Scores:  
 Pred. No.: 1.5e-104 Length: 249  
 Score: 1387.00 Matches: 248  
 Percent Similarity: 99.60% Conservative: 0  
 Best Local Similarity: 99.60% Mismatches: 1  
 Query Match: 95.13% Indels: 0  
 DB: 4 Gaps: 0  
 US-09-598-982-20 (1-771) x US-09-079-970A-5 (1-249)  
 QY 7 CTCGAGAAAAAGAACATGGTGGGGTCAGGAGGCCAGGAGCAGTGSCCCCTGGCAGGT 66  
 Db 1 LeuGluLysArgIleValGlyGlyGlnGluAlaProArgSerLysTrpProtGlnVal 20  
 QY 67 AGCCGTGAGAAGTCCACGGGCCATACTGGAGCACUTCTCGGGGGCTCCCTCATCCACCC 126  
 Db 21 SerLeuArgValHisGlyProTrpMetThisPheCysGlyGlySerLeuIleHisPro 40  
 QY 127 CAGTGGGGCTGACCCGGGCTCGTGGGACCGGACTCAAGATCTGGCCCTC 186  
 Db 41 GluTrpValIleThrIleAlaHisCysValGlyProAspPheValAlaLeu 60  
 QY 187 AGGGTGAACACTGGGGAGAGCACCTCTACTACAGGACAGCTGTCGGGTCAAGG 246  
 Db 61 ArgValGlnIleArgGluGlnHisLeuTyroGlnAspGlnIleLeuProValSerArg 80  
 QY 247 ATCATCGTCACCCACAGTCTCACCCCAATATGGAGCAGCATTGGAGCAGATGCCCTGCTSG 306  
 Db 81 IleLeuIleValHisProGlnIlePheIleThrAlaGlnIleGlyAlaAspIleAlaLeuGlu 100  
 QY 307 CTGGAGGAGCCGAGGAACTCTCCAGCCACGTCACAGGTACCCCTGGCCCTCA 366  
 Db 101 LeuGluIleIleProLysValSerSerHisValIleThrLeuProAlaLaser 120  
 QY 367 GAGACTCTCCCCGGGGTGCCTCGCTCACTGGCTGGGGCATCTGGACATGAT 426  
 Db 121 GluThrPheProProGlyMetProCysTrpValIleGlyTrpGlyIleAspValAspAsn 140  
 QY 427 GAGCCCTTCACCGCCATTCTGAGCAGGTGAAGCTCCCATAATGGAAACAC 486  
 Db 141 GluArgLeuProProPheProPheProLeuLysGlnValIleValProIleMetGluAsnHis 160  
 QY 487 ATTCTGTGACGCCAAATACACCTTGCCCTCACGGGAGACCCGCGCCGCTGCTGGCCTG 546  
 Db 161 IleCysAspAlaIleSthIleLeuIleAlaTyrrHrgIleAspPheValArgIleValArg 180  
 QY 547 GACGACATGCTGCTGCCGACACCCGGAGGGACTCATGCCAGGGGACTCGGAGGG 606  
 Db 181 AspAspMetIleCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGly 200  
 QY 607 CCCCGGTTGCAAGGTGATGGCACCTGGCTGCTGAGGGGGCTGAGCTGGCCAG 666  
 Db 201 ProLeuValCysLysValAsnGlyIleTrpLeuGlnAlaGlyValValSerTrpGlyGlu 220  
 QY 667 GGCGTGGCCAGGCCAACGGCCCTGGCAGCTACACCCGGTCACCTACTGGACTGG 726  
 Db 221 GlyCysAlaGlnProAsnArgProGlyIleTyrrHrgValIleThrTyrrIleLeuAspTrp 240  
 QY 727 ATCCACCACTATGCCCAAACCCG 753  
 Db 241 IleIleHistIleIleValProLysLysPro 249  
 RESULT 2  
 US-09-079-970A-6  
 ; Sequence 6, Application US/09/079/970A  
 ; Patent No. 6,274,366  
 ; GENERAL INFORMATION:

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/079, 970A  
 REFERENCE/DOCKET NUMBER: 34506. 073  
 FILING DATE:  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Leone, Joseph T.  
 REGISTRATION NUMBER: 37, 170  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (608) 831-2106  
 TELEFAX: (608) 831-2106  
 INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 245 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-09-079-970A-6

Alignment Scores:  
 Pred. No.: 5.12e-103 Length: 245  
 Score: 1368.00 Matches: 244  
 Percent Similarity: 99.59% Conservative: 0  
 Best Local Similarity: 99.59% Mismatches: 1  
 Query Match: 93.83% Indels: 0  
 DB: 4 Gaps: 0  
 US-09-598-982-20 (1-771) x US-09-079-970A-6 (1-245)  
 QY 19 ATCGTGGGGCTAGGAGCCCCAGGAGCAAGTGGCCCTGGAGGTGAGGAGTC 78  
 Db 1 IleValGlyGlyGlnGluAlaProArgSerLysTrpProtGlnValSerLeuArgVal 20  
 QY 79 CACGCCCTACTGATGACTCTTGCGGGGCTCCATCCACCCAGGGCTG 138  
 Db 21 HisGlyProTrpMetThisPheCysGlyGlySerLeuIleHisProGlnIleVal 40  
 QY 139 ACCGGCGGGCTGGTGGGAGGCTCAAGGATCTGGCCCTCAGGGTCAC TG 198  
 Db 41 ThrAlaAlaIleHisCysValGlyProAspValAlaIleLeuArgIleGlu 60  
 QY 199 CGGAGCACCCCTACTACAGGACAGCTGCTGCGGTAGCAGGATCATGTCAC 258  
 Db 61 ArgGluGlnHisLeuTyroGlnAspGlnIleLeuProValSerArgIleLeuHis 80  
 QY 259 CCACAGTCTACACGCCAGTCGGAGGGAGATCGGCTCTGGACTGAGGAGCCG 318  
 Db 81 ProGlnIlePheIleAlaIleIleGlyAlaAspIleAlaLeuLeuGluIleGlu 100  
 QY 319 GTGAGGTCCTCACGCCACCTCACGGTCACCCCTGGCCCTCAGAGCCTTCCC 378  
 Db 101 ValIleValSerSerHisValIleThrIleLeuProProAlaSerGluIlePhePro 120  
 QY 379 CGGGGATGCCGCTGGCTCACGAGCTGGCTGGGAGGACATGATGAGGCCCTCCA 438

Db 121 ProGlyMetProCysTrpValThrCysTyrPglYaspValAspAsnAspGluArgLeuPro 140  
 Qy 439 CCGCCATTCTTCAGGAGGTAAAGGTCATGCCCTAATGAAACACATTGGACCA 498  
 Db 141 ProPropheproleuysGlnValIysValProIemMetGluasnHistilecysAspAla 160  
 Qy 499 AAATCCACCTGGCCCTACACGGAGCACGACCTCCGCATCGCCGAGCAGCTG 558  
 ||||||| 161 LysThrHisLeuGlyAlaTyrThrGlyaspAspValArgLeuValGlyProLeu 180  
 Qy 559 TGTGGGGACACCGGGAGGACTCATGCCAGGGACTCCGGGAGGGCCCTGGTG 618  
 181 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeu 200  
 Qy 619 AAGGGAATGCACTGGTGAGCGGGCGCTGGTAGTTGGGGAGGGCTG 678  
 ||||||| 201 LysValAsnGlyThrTyrPheUglnAlaGlyValValSerTrpGlyGlu 220  
 Qy 679 CCCACCCGGCTGCATCACCGGTCACACTTGACTGGATGGATCCACCAT 738  
 ||||||| 221 ProAsnArgProGlyIleTyrTyrArgValThrTyrTyruAspTrpIleH 240  
 Qy 739 GTCCCCAAAGCCG 753  
 |||||||  
 Db 241 ValProLysPro 245  
 RESULT 3  
 US-09-016-366A-21  
 ; Sequence 21, Application US/09016366A  
 ; Patent No. 5955431  
 GENERAL INFORMATION:  
 APPLICANT: Stevens, Richard L.  
 TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE  
 TITLE OF INVENTION: INHIBITORS  
 NUMBER OF SEQUENCES: 65  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
 STREET: 600 Atlantic Avenue  
 CITY: Boston  
 STATE: MA  
 COUNTRY: U.S.A.  
 ZIP: 02210-2211  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Disquette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSEED for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/016, 366A  
 FILING DATE: January 30, 1998  
 CLASSIFICATION: 530  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 60/037, 090  
 FILING DATE: 05-FEB-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Plumer, Elizabeth R.  
 REGISTRATION NUMBER: 36, 637  
 REFERENCE/DOCKET NUMBER: B0801/7093  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-720-3500  
 TELEX: 617-720-2441  
 INFORMATION FOR SEQ ID NO: 21:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 274 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 US-09-016-366A-21  
 Alignment Scores:  
 Pred. No.: 5.25e-103 Length: 274

Score: 1368 00 Matches: 244  
 Percent Similarity: 99.59% Conservative: 0  
 Best Local Similarity: 99.59% Mismatches: 1  
 Query Match: 93.83% Indels: 0  
 Gaps: 2

US-09-598-982-20 (1-771) x US-09-016-366A-21 (1-274)

Qy 19 ATCGTGGGGGGTCAAGGAGGCCAGTGCCCTGCGAGGTGAGCAG 78  
 ||||||| 30 IleValGlyGlyGlnAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 49  
 Qy 79 CACGGCCCATACTGGATGCACTCTGGGGGCTCCATCACCCCACTGGGGCTG 138  
 ||||||| 50 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProDlnTrpValLeu 69  
 Qy 139 ACCGGCGGGGGTGGGGGAGCGAGTCAGGATCTGGCCCTCAGGGCACTG 198  
 ||||||| 70 ThrAlaAlaHisCysValGlyProAspValAspLeuAlaAlaLeuArgValGlnLeu 89  
 Qy 199 CGGAGGACACCTCTACACAGGACGACCAGCTGCGCGTACAGGATCArgGTG 258  
 ||||||| 90 ArgGluGlnHisLeuIrryGlyGlnAspGlnLeuLeuProValSerArgIleIleValHis 109  
 Qy 259 CCACAGTCTCACCCGCCAGATCGGCCCTCTGGAGCTGGAGTGGAGGGCC 318  
 ||||||| 110 ProGlyInPhenTyrThrAlaGinIleGlyAlaAspIleAlaLeuLeu 129  
 Qy 319 GTGAAGGTCTCAGGACAGTCACAGGTACACGGTACCCGGCCCTGCTCAGAGACCTTCCC 378  
 ||||||| 130 ValIysValSerSerHistValHistValThrLeuProProAlaLaserGluThrPhePro 149  
 Qy 379 CCGGGATGCGCTGGGRCAGTGCCTGGCCATGPGACAAATGATGAGGAGCCTCC 438  
 ||||||| 150 ProGlyMetProCysTrpValThrGlyTyrPglYaspValAspAsnAspGluArgLeuPro 169  
 Qy 439 CCGCCATTCTTGAGCAGGTGAGGTCCCATATGAAACACATTGGACCA 498  
 ||||||| 170 ProPropheproleuysGlnValIysValProIemMetGluasnHistilecysAspAla 189  
 Qy 499 AAATCCACCTGGCCCTACACGGAGACGACGGTCCGATCGRCGAGCAGACATG 558  
 ||||||| 190 LysThrHisLeuGlyAlaTyrThrGlyaspAspValArgLeuValGlyProLeu 209  
 Qy 559 TGTGGGGACACCGGGAGGACTCATGCCAGGAGCTCCGGAGGSCCCCTGGTG 618  
 ||||||| 210 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeu 229  
 Qy 619 AAGGGAATGCACTGGTGAGCGGGCGCTGGCAGCTGGGGAGGGCTGGCC 678  
 ||||||| 230 LysValAsnGlyThrTyrPheUglnAlaGlyValValSerTrpGlyGlu 249  
 Qy 679 CCCACCCGGCTGGGATCACACCGGTGTCACACTTGACTGGATGTCACCACT 738  
 ||||||| 250 ProAsnArgProGlyIleTyrThrArgValThrTyrTyruAspTrpIleH 269  
 Qy 739 GTCCCCAAAGCCG 753  
 ||||||| 270 ValProLysPro 274

RESULT 4  
 US-08-978-404B-16  
 ; Sequence 16, Application US/08978404B  
 ; Patent No. 5968782  
 GENERAL INFORMATION:  
 APPLICANT: Stevens, Richard L.  
 TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES  
 TITLE OF INVENTION: FIBRINOGEN  
 NUMBER OF SEQUENCES: 74  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
 STREET: 600 Atlantic Avenue  
 CITY: Boston

STATE: MA  
 COUNTRY: U.S.A.  
 ZIP: 02210-2211  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSeq for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/978, 404B  
 FILING DATE: 25-NOV-97  
 CLASSIFICATION: 435  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 60/032, 354  
 FILING DATE: 04-DEC-1996  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Plumer, Elizabeth R.  
 REGISTRATION NUMBER: 36, 637  
 REFERENCE/DOCKET NUMBER: B0801/7090  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-720-3500  
 TELEFAX: 617-720-2441  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 16:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 274 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: No. 5968782e  
 US-08-978-404B-16

Alignment Scores:  
 Pred. No.: 5.25e-103   Length: 274  
 Score: 1368.00   Matches: 244  
 Percent Similarity: 99.59%   Conservative: 0  
 Best Local Similarity: 99.59%  
 Query Match: 93.83%  
 DB: 2

US-09-598-982-20 (1-771) x US-08-978-404B-16 (1-274)

US-09-016-366A-19 ; Sequence 19, Application US/09016366A  
; Patent No. 595531  
; GENERAL INFORMATION:  
; APPLICANT: Stevens, Richard L.  
; APPLICANT: Huang, Chifu  
; TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE  
; NUMBER OF SEQUENCES: 65  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
; STREET: 600 Atlantic Avenue  
; CITY: Boston  
; STATE: MA  
; COUNTRY: U.S.A.  
; ZIP: 02210-2211  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/016, 366A  
; FILING DATE: January 30, 1998  
; CLASSIFICATION: 530  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: 60/037, 090  
; FILING DATE: 10-FEB-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Plumer, Elizabeth R.  
; REGISTRATION NUMBER: 36, 637  
; REFERENCE/DOCKET NUMBER: B0801/7093  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-720-3500  
; TELEFAX: 617-720-2441  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 273 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-09-016-366A-19

Alignment Scores:  
 Pred. No.: 1.33e-102   Length: 273  
 Score: 1363.00   Matches: 243  
 Percent Similarity: 99.18%  
 Best Local Similarity: 99.18%  
 Query Match: 93.48%  
 DB: 2

QY 499 AAATPACCCTTGGCCTTACACGGAGCACGACTCCGTGACCATGCTG 558  
 |||||||CCTTGGCCTTACACGGAGCACGACTCCGTGACCATGCTG 558  
 Db 190 LYSTYRHISLeuGlyAlaTyrrhinglyAspaspValArgleuValargaspaspMetIeu 209  
 QY 559 TGTGCCGAAACCCGAGGGACTCATGCCGGACTCGGGAGGGCCCTGGGTC 618  
 |||||||CCTTGGCCTTACACGGAGCACGACTCCGTGACCATGCTG 618  
 Db 210 CysAlaGlyAsnThrArgArgAspSerCysGlyAlaValSerTrpGlyGluGlyAsnAlaGln 229  
 QY 619 AGGGTGAATGGCACCTGGCTGAGGGGGGGTGAGCTGGGGCTGTGCCAG 678  
 |||||||CCTTGGCCTTACACGGAGCACGACTCCGTGACCATGCTG 678  
 Db 230 LystaLAslGlyThrTrpIeuGlyAlaValSerTrpGlyGluGlyAsnAlaGln 249  
 QY 679 CCCAACCGGCCCTGGCATCTACACCGGTGTCACCTACTACTGACTGATCCACCATAT 738  
 |||||||CCTTGGCCTTACACGGAGCACGACTCCGTGACCATGCTG 738  
 Db 250 ProAsnArgProlGlyIleTyrrhargValThrTyrrhLeuAspTrpIleHishistYr 269  
 QY 739 GTCCCCAAAAGGG 753  
 |||||||CCTTGGCCTTACACGGAGCACGACTCCGTGACCATGCTG 753  
 Db 270 ValProLysLysPro 274

RESULT 5  
 US-09-016-366A-19 ; Sequence 19, Application US/09016366A  
; Patent No. 595531  
; GENERAL INFORMATION:  
; APPLICANT: Stevens, Richard L.  
; APPLICANT: Huang, Chifu  
; TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE  
; NUMBER OF SEQUENCES: 65  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
; STREET: 600 Atlantic Avenue  
; CITY: Boston  
; STATE: MA  
; COUNTRY: U.S.A.  
; ZIP: 02210-2211  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/016, 366A  
; FILING DATE: January 30, 1998  
; CLASSIFICATION: 530  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: 60/037, 090  
; FILING DATE: 10-FEB-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Plumer, Elizabeth R.  
; REGISTRATION NUMBER: 36, 637  
; REFERENCE/DOCKET NUMBER: B0801/7093  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-720-3500  
; TELEFAX: 617-720-2441  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 273 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-09-016-366A-19

QY 499 AAATPACCCTTGGCCTTACACGGAGCACGACTCCGTGACCATGCTG 558  
 |||||||CCTTGGCCTTACACGGAGCACGACTCCGTGACCATGCTG 558  
 Db 190 LYSTYRHISLeuGlyAlaTyrrhinglyAspaspValArgleuValargaspaspMetIeu 209  
 QY 559 TGTGCCGAAACCCGAGGGACTCATGCCGGACTCGGGAGGGCCCTGGGTC 618  
 |||||||CCTTGGCCTTACACGGAGCACGACTCCGTGACCATGCTG 618  
 Db 210 CysAlaGlyAsnThrArgArgAspSerCysGlyAlaValSerTrpGlyGluGlyAsnAlaGln 229  
 QY 619 AGGGTGAATGGCACCTGGCTGAGGGGGGGTGAGCTGGGGCTGTGCCAG 678  
 |||||||CCTTGGCCTTACACGGAGCACGACTCCGTGACCATGCTG 678  
 Db 230 LystaLAslGlyThrTrpIeuGlyAlaValSerTrpGlyGluGlyAsnAlaGln 249  
 QY 679 CCCAACCGGCCCTGGCATCTACACCGGTGTCACCTACTACTGACTGATCCACCATAT 738  
 |||||||CCTTGGCCTTACACGGAGCACGACTCCGTGACCATGCTG 738  
 Db 250 ProAsnArgProlGlyIleTyrrhargValThrTyrrhLeuAspTrpIleHishistYr 269  
 QY 739 GTCCCCAAAAGGG 753  
 |||||||CCTTGGCCTTACACGGAGCACGACTCCGTGACCATGCTG 753  
 Db 270 ValProLysLysPro 274

US-09-598-982-20 (1-771) x US-09-016-366A-19 (1-273)

QY 19 ATGGTCGGGTCAGAGGCCAGAGCACTGGCCCTGCAGTGAGCTGAGTC 78

Db 29 IleValGlyIglnGluAlaProArgSerIleValProGlnValIleArgVal 48

QY 79 CACGCCCATACTGGATGCACTCTGGGGGCTCCCTCATCACCCCCAGTGGGCTG 138

Db 49 HisGlyProTryrPmrPheCysLyGlySerIleIleHisProGlnIleProVal 68

QY 139 ACCGCCCGGGCTGGGACCGGCTCACGGATCTGGCCCTCACGGTGACTG 198

Db 69 ThrAlaLahisCysValGlyProAspValLeuAlaAlaLeuArgValInLeu 88

Db 89 ArgGluGlnHistLeuIleUtyrGlnAspIleValLeuProValSerArgIleIleValHis 108

QY 199 CGGAGGAGCACCTCTACTACAGGAGCCAGTCGAGGACATCGGCTGAGCTGAC 258

Db 109 ProGlnPhyterPheIleGlnIleGlyAlaAspIleAlaLeuGluIleGluIlePro 128

QY 319 GTCAGGTTCTCCAGCTCACGGTCCACAGTCACGGTCTGGCCCTGCTGAGCTCCC 378

Db 129 ValAsnValSerSerHisValHistValThrLeuIleProProAlaSerLaserGluThrPhePro 148

QY 379 CCGGGGAGGCCGCGCTGGTACTGGCCGGGAGATGGAAACACATTTGAGCC 318

Db 149 ProGlyMetProCysTrpValIleIleGlyIleValProIleMetGluAsnHisIleCysAspAla 168

QY 439 CCCGCAATTCTCTGAAAGGCTGAGGTCCCATAATGGAAACACATTTGAGCC 498

Db 169 ProProPheProLeuIleValProIleMetGluAsnHisIleCysAspAla 188

QY 499 AAATACCAACCTGGCCCTACAGGGAGACAGCTGGCCATGTCCTGGACATCTG 558

Db 189 LysTyrHisLeuGlyAlaTyrrGlyAspAspValArgIleValArgGluAspPheMetIeu 208

QY 559 TGGCCTGGAACACCCTGGAGGACTCAGGCCAGGGCAGCTCGGGGAGGGCTGGCTG 618

Db 209 CysAlaGlyAsnThrArgAspSerCysGlnGlyProLeuValIleValHis 228

QY 619 AACGTGAATGGCACCTGGCTGAGGGGGCTGGTACCTGGGGAGGGCTGGCC 678

Db 229 LysValAsnGlyIleThrPheIleGlnIleGlyValAspIleTrpGlyGluIleCysAlaGln 248

QY 679 CCCAACCGGCCCTGGCTACACCCGTCACCTACTCTGGACTGATGCCACCAT 738

Db 249 ProAsnArgProGlyIleTyrrThrArgValThrTyrIleLeuAspTrpIleHisIleTyr 268

QY 739 GRCGCCAAAGGCC 753

Db 269 ValProIleGlyPro 273

RESULT 6  
US-09-978-404B-14

; Sequence 14, Application US/08978404B  
; Patent No. 5,969,782  
; GENERAL INFORMATION:

; APPLICANT: Stevens, Richard L.  
; TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES  
; TITLE OF INVENTION: FIBRINOGEN  
; NUMBER OF SEQUENCES: 74  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.

; STREET: 600 Atlantic Avenue  
; CITY: Boston  
; STATE: MA  
; COUNTRY: U.S.A.  
; ZIP: 02110-2211

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS

SOFTWARE: PassSeq for Windows Version 2.0

CURRENT APPLICATION DATA: For Windows Version 2.0

APPLICATION NUMBER: US/09-978, 404B

FILING DATE: 25-NOV-97

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/032, 354

FILING DATE: 04-DEC-1995

ATTORNEY/AGENT INFORMATION:

NAME: Plummer, Elizabeth R.

REGISTRATION NUMBER: 36-637

REFERENCE/DOCKET NUMBER: B0801/7090

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-720-3500

TELEFAX: 617-720-2441

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 273 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: NO. 5968782e

US-09-598-982-20 (1-771) x US-09-978-404B-14 (1-273)

Alignment Scores:

Score: 1.33e-102

Length: 273

Percent Similarity: 99.18%

Best Local Similarity: 99.18%

Query Match: 93.48%

DB: 2

Gaps: 0

US-09-598-982-20 (1-771) x US-09-978-404B-14 (1-273)

QY 19 ATGGTCGGGGTAGGAGGCCAGAGCACTGGCCCTGCAGGGAGCC 78

Db 29 IleValGlyIglnGluAlaProArgSerIleValProGlnValIleArgVal 48

QY 79 CACGCCCATACTGGATGCACTCTGGGGGCTCCCTCATCACCCCAGGGCTG 138

Db 49 HisGlyProTryrPmrPheCysLyGlySerIleIleHisProGlnIleProVal 68

QY 139 ACCGCCGGGCGGGAGGCTCAAGGATCTGGCCCTCAGGGCTGAGCTG 198

Db 69 ThrAlaLahisCysValGlyProAspValLeuAlaAlaLeuArgValInLeu 88

QY 199 CGGGAGCAGCACCTACTACAGGAGGCTCTGGGGAGATGAGGGCTCCC 258

Db 89 ArgGluGlnHistLeuIleUtyrGlnAspIleValLeuProValSerArgIleIleValHis 108

QY 259 CACAGTTCTACACGCCAGATGGAGCCATCCCTGAGGGCTGAGCTG 318

Db 109 ProGlnPhyterPheIleValProIleMetGluAsnHisIleCysAspAla 128

QY 319 GTGGAGGCTCAGGCCAGTCACACGTCACCTGGGGGATGTGGAGGGCTCCC 378

Db 129 ValAsnValSerSerHisValHistValProIleGluIleLeuGluIlePro 148

QY 379 CCGGGATGCGGCTCTGGCTACTGGCTGGGGATGTGGAGGGCTCCC 438

Db 149 ProGlyMetProCysTrpValIleIleGlyIleValProAspIleAspIlePro 168

QY 439 CCCCATTCCTCTGAACCGAGGTAAGGATCTGGCCCTCAGGGCTGAGCTG 498

Db 169 ProProPheProLeuIleValProIleMetGluAsnHisIleCysAspAla 188

QY 499 AAATACCAACCTGGCTACACGGAGCAGCTGGCTGGGAGACAGCTG 558

Db 189 LysTyrHisLeuGlyAlaTyrrGlyAspAspValArgIleValGluAspPheMetIeu 208

QY 559 TGTGCCGGGACCCGGAGGACTCATGCCAGGGCACTCGGAAGGGCCCTGGTGTG 618  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 209 CYSAlaGlyAsnThrArgAspSerCysGlnGlyAspSerGlyProLeuValCys 228  
 QY 619 AAGGTGAATGGCACCTGGCTGCACTACTGAGCTGAGTCACCAT 738  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 249 ProAsnArgProGlyIleTyrrhArgValThrTyrrhLeuAspPileHishTyr 268  
 QY 739 GTCGCCAAAGCCG 753  
 ||||| ||||| |||||  
 Db 269 ValProLysPro 273

RESULT 7

Sequence 23, Application US/09016366A  
 Patent No. 5955431  
 GENERAL INFORMATION:  
 APPLICANT: Stevens, Richard L.  
 TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE  
 NUMBER OF SEQUENCES: 65  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
 STREET: 600 Atlantic Avenue  
 CITY: Boston  
 STATE: MA  
 COUNTRY: U.S.A.  
 ZIP: 02210-2211  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSEQ for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/016,366A  
 FILING DATE: January 30, 1998  
 CLASSIFICATION: 530  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 60/037, 090  
 FILING DATE: 05-FEB-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Clumer, Elizabeth R.  
 REGISTRATION NUMBER: 36,637  
 REFERENCE DOCKET NUMBER: B0801/7093  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-720-3500  
 TELEX:  
 TELEFAX: 617-720-2441

INFORMATION FOR SEQ ID NO: 23:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 267 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear

US-09-016-366A-23

Alignment Scores:  
 Pred. No.: 4.54e-101  
 Score: 1.44.00  
 Percent Similarity: 98.37%  
 Best Local Similarity: 98.37%  
 Query Match: 92.18%  
 DB: 2  
 Gaps: 0

US-09-598-982-20 (1-771) x US-09-016-366A-23 (1-267)

QY 19 ATCGTGGGGTCAGGAGCCCCAGGCAAGTGCCCTGAGCTGAGAGTC 78  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 23 IleValGlyGlyInGluIaProArgSerIysTrPProtPglnValSerLeuArgVal 42  
 QY 79 CACGGCCATACATGGGACTCTCTGGGGGCTCCTCATCACCCCAGTGGGCTG 138  
 Db 43 ArgAspArgTyrrhPmetHisPheCysGlyGlySerIleLeuHishProGlnPvalLeu 62  
 QY 139 ACGGCCGGCTGGGGACGGACGGCTAGAGCTGGCCCTCAGGGCACTGGGCACTG 198  
 ||||| |||||  
 Db 63 ThrAlaLahIisCysValGlyProAspValAspIleAlaLeuArgValGlnLeu 82  
 QY 199 CGGAGGAGCACCTACTACAGAGGACAGCITGCTGCCGTCAGCAGGATCATCGTC 258  
 ||||| |||||  
 Db 83 ArgGluGlnHisLeuTyrrhGlnAspGlnIleLeuProValSerArgIleLeuValHis 102  
 QY 259 CGACAGSTCTACACC5GCCAGATCGGCTGAGCAGTCAGCAGGAGCC 318  
 ||||| |||||  
 Db 103 ProGlnPhetYrrhAlaGlnIleGlyAlaAspIleAlaLeuLeuGluGlu 122  
 QY 319 GCGAAGCTCTCCAGCAGTCCACAGCTCACCTGCCCTGCCCTCAGGACCTTCCC 378  
 ||||| |||||  
 Db 123 ValIysValSerHisValHishTyrValIleLeuProProlaserGluThrPhePro 142  
 QY 379 CGGGGGATGCGTGCAGGGTACTGGCTGGGGCATGTGGAATGATGAGCCCTCCA 438  
 ||||| |||||  
 Db 143 ProGlyMetProCystItpValThrGlyTrpGlyAspValAspAspGluArgLeuPro 162  
 ||||| |||||  
 Db 439 CGGCCATTTCCCTCTGAACAGGTGAAGTGTGCCATATGAAAACACATWTGAGCA 498  
 ||||| |||||  
 Db 163 ProProHeProIleUlysGlnValIysValProIleMetGluAsnHishIleCysAspAla 182  
 QY 499 AsATACCACCTGGCCTACCGGGAGACRGTCGGCATGTCGTCGTCGAGACATGCTG 558  
 ||||| |||||  
 Db 183 LysTyrIleLeuGlyIleTyrIleLeuIleLeuIleLeuIleLeuIleLeuIleLeu 202  
 QY 559 TGTGCCGGAACCCGGAGGGACTCATGCCAGGGACTCCGGAGGGCCCTGGTGC 618  
 ||||| |||||  
 Db 203 CysAlaGlyAsnThrArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 222  
 QY 619 AAGGTGAATGGCACCTGGCTGAGGCGGGCTGGTCAGCTGGGGAGGGCGTGC 678  
 ||||| |||||  
 Db 223 LysValAsnGlyIleTyrrhArgValThrTyrrhLeuAspPileHishTyr 242  
 QY 679 CCAACGGGCCAGCCTACCCGCGTCACTACTGAGCTGAGTCACCACT 738  
 ||||| |||||  
 Db 243 ProAsnArgProGlyIleTyrrhArgValThrTyrrhLeuAspPileHishTyr 262  
 QY 739 GTCGCCAAAGCCG 753  
 ||||| |||||  
 Db 263 ValProLysPro 267

RESULT 8

US-08-978-404B-18

Sequence 18, Application US/08978404B  
 Patent No. 5968782  
 GENERAL INFORMATION:  
 APPLICANT: Stevens, Richard L.  
 TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES  
 NUMBER OF SEQUENCES: 74  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
 STREET: 600 Atlantic Avenue  
 CITY: Boston  
 STATE: MA  
 COUNTRY: U.S.A.  
 ZIP: 02210-2211  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSEQ for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/978,404B

FILING DATE: 25-NOV-97  
 CLASSIFICATION: 435  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 60/032,354  
 FILING DATE: 04-DEC-1996  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Plumer, Elizabeth R.  
 REGISTRATION NUMBER: 36,637  
 REFERENCE/DOCKET NUMBER: B0801/7090  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-720-3500  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 18:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 267 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: NO. 5968782e  
 ; US-08-978-404B-18  
 Alignment Scores:  
 Pred. No.: 4.54e-101 Length: 267  
 Score: 1344.00 Matches: 241  
 Percent Similarity: 98.37% Conservative: 0  
 Best Local Similarity: 98.37% Mismatches: 4  
 Query Match: 92.18% Indels: 0  
 DB: 2 Gaps: 0  
 US-09-598-982-20 (1-771) x US-08-978-404B-18 (1-267)  
 QY 19 ATCGTGGGGTCAGAGGAGGCCAGAGTGCCCTGGCAGGTGACCTGAGTC 78  
 Db 23 IleValGlyGlyGlnGluIalaproArgSerLysTrpProTrpGinValSerLeuArgVal 42  
 QY 79 CACGCCCATACTSGATGCACTTCGGGGCTCCTCATCCACCCCCAGTGGGCTG 138  
 Db 43 ArgAspArgTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGinTrpValLeu 62  
 QY 139 ACCGGCGGGGGTGGGGACGGGACGTCAGAGGCTGGCCCTAGGTGGACTG 198  
 Db 63 ThraAlaAlaHisCysValGlyProSpvalLysPleuAlaAlaLeuIleArgVal 82  
 QY 199 CGGGAGCAGCACCTGACTACCAGGACACTGCGGCCGTCAGGAGCATGTCAC 258  
 Db 83 ArgGluGlnHisLeuTyrryrglnAspGinLeuLeuuproValSerArgIleLeuValHis 102  
 QY 259 CCACAGTTCPACACGCCAGATCGAGGAGCACGCCCTGCTGAGCTGGACGCC 318  
 Db 103 ProGlnPhyTyrThrAlaGinIleGlyAlaAspPileAlaLeuLeuIleGluIleLeu 122  
 QY 319 GTGAAGGTCTCAGCACGTCACAGGTGACCCCTGCCCTGCTCAGAGCTCCC 378  
 Db 123 ValysValserSerHisValHisThrValThrLeuProAlaSerGluIlePhPro 142  
 QY 379 CCGGCATGCGCTGCGGGACTCGTGGGAGCAAGATGAGCCCTCCCA 438  
 Db 143 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAspGluIleLeuPro 162  
 QY 439 CCGGCATGCGCTGCGGGACTCGTGGGAGCAAGATGAGCCCTCCCA 498  
 Db 163 ProPropheProLeuIysGinValIysValProIleMetGluIshisIleCysAla 182  
 QY 499 AAATACACCTGGGCCAACGGAGAGCACGCCATGAGCTGACGACATGCTG 558  
 Db 183 LysTrpHistLeuGlyAlaIleTrpIleLeuArgAspAspMetLeu 202  
 QY 559 TGTGGCGGGACACCGGAGGGACTCATGCGAGGCCGACTCCGGGGGCCCTGTC 618  
 Db 203 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 222  
 QY 619 AAGGTTGAATGGCACCTGGCTGCAGCGGGGGCTGCTGGGGAGGGCTGGCCAG 678

---

RESULT 9  
 US-08-944-483-69  
 Sequence 69, Application US/08944483  
 ; Patent No. 6232456  
 GENERAL INFORMATION:  
 ; APPLICANT: COLETTI, TRACEY L.  
 ; APPLICANT: FRIEDMAN, PAULA N.  
 ; APPLICANT: GRANADOS, EDWARD N.  
 ; APPLICANT: KLAAS, MICHAEL R.  
 ; APPLICANT: ROUSSELL, JOHN C.  
 ; APPLICANT: STEWART, KENT D.  
 ; APPLICANT: STROUSE, STEVEN D.  
 TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS  
 TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES  
 TITLE OF INVENTION: OF THE PROSTATE  
 NUMBER OF SEQUENCES: 76  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Abbott Laboratories  
 STREET: 100 Abbott Park Road  
 CITY: Abbott Park  
 STATE: IL  
 COUNTRY: USA  
 ZIP: 60064-3500  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FASTSEQ for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08-944-483  
 FILING DATE:  
 CLASSIFICATION: 424  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Becker, Cheryl L.  
 REGISTRATION NUMBER: 35,441  
 REFERENCE/DOCKET NUMBER: 6183.US.01  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 847/938-2623  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 69:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 245 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: NO. 6232456  
 Alignment Scores:  
 Pred. No.: 2.02e-95 Length: 245  
 Score: 1274.00 Matches: 226  
 Percent Similarity: 94.69% Conservative: 6  
 Best Local Similarity: 92.24% Mismatches: 13  
 Query Match: 87.38% Indels: 0  
 DB: 4 Gaps: 0  
 US-09-598-982-20 (1-771) x US-08-944-483-69 (1-245)







QY 118 ATCCACCCCCCAGGGGCTCTGACCGCCCGGCTGGGGACGGTCAAGGATCTG 177  
 QY ||||||| ||||||| ||||||| ||||||| ||||||| : : : : :  
 Db 65 IleHisProGlnTrpValLeuThrAlaLalaIscysValGlyProHisIleSerPro 84  
 QY 178 GCGCCCTCAGGGGCAACTGGGGGACAGCACCTCTACTACAGGACCGAGCTGCG 237  
 Db 85 GlnLeuPheArgValGlnLeuArgLysGluIntrpLeuTyrrTyrglyAspGlnLeuSer 104  
 QY 238 GTCAGCAGGATCATCGTCACCCACAGCTACACGCCAGTCGGACGGACAGCTGCG 297  
 Db 105 LeuAsnAspGlyValProHisProIleSerTrpValIleGluGlyAla 124  
 QY 358 CCTGCCTCAGAGACCTTCCCCGGGGTGCCTGCGGGTCACTGGGCTACTGGGCTG 417  
 Db 145 ProLaserGluThrPheProProGlyThrSerCysTrpValThrGlyLysPro 164  
 Db 125 LeuLeuGluLeuGluValProValAsnValSerThrHisLeuHisProIleSerLeuPro 144  
 QY 478 GAAACACATTTGTGACCAAATACACACCTTGCGCTACACGGGAGCACGAGCTCGC 537  
 Db 185 GluAsnSerLeuCysAspArgLysTrpIleGlyMetLeuCysAlaGlyAsp 204  
 QY 538 ATCGTCGCGGACGQATCTGTCGCGGGAGACCCGGAGGGACTCTGCCAGGCCAC 597  
 Db 205 IleValHisAspGlyMetLeuCysAlaGlyAspSerCysGlnGlyAsp 224  
 QY 598 TCCGAGGCCCTGGTCAAGGGTGAATGGCACTGGTGGAGGGGGCGGTGCG 657  
 Db 225 SerGlyGlyProLeuValProLeuProProGlyThrTrpLeuGlnAlaGlyValAsp 184  
 QY 658 TGGGGAGGGCTGTGCCAGCCACCCGCTGCGATCTACACCGTGTACACTAC 717  
 Db 245 TrpGlyLysGlyCysAlaGlnProAsnLysProGlyLysLeuArgValThrTyr 264  
 QY 718 TTGGACTGATCCACCACTATGTCCTCAA 747  
 Db 265 LeuAspTrpIleHisArgTyrValProGlu 274  
 RESULT 14  
 US-08-978-404B-8  
 ; Sequence 8, Application US/08978404B  
 ; General Information:  
 ; APPLICANT: Richard L.  
 ; TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES  
 ; NUMBER OF SEQUENCES: 74  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
 ; STREET: 600 Atlantic Avenue  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: U.S.A.  
 ; ZIP: 02210-2211  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastRC for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/978,404B  
 ; FILING DATE: 25-NOV-97  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 60/032,354  
 ; FILING DATE: 04-DEC-1996

ATTORNEY/AGENT INFORMATION:  
 NAME: Plumer, Elizabeth R.  
 REGISTRATION NUMBER: 36 637  
 REFERENCE/DOCKET NUMBER: B0801/7090  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-720-3500  
 TELEFAX: 617-720-2441  
 INFORMATION FOR SEQ ID NO: 8:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 270 amino acids  
 TOPOLOGY: linear  
 STRANDEDNESS: single  
 TYPE: amino acid  
 TOPOL: NO. 5988782e  
 US-08-978-404B-8

Alignment Scores:  
 Pred. No.: 2.48e-79 Length: 270  
 Score: 1075.00 Matches: 188  
 Percent Similarity: 84.7% Conservative: 18  
 Best Local Similarity: 77.3% Mismatches: 37  
 Query Match: 73.7% Indels: 0  
 DB: 2 Gaps: 0

US-09-598-982-20 (1-771) x US-08-978-404B-8 (1-270)

QY 19 ATCCCTGGGGCTGGAGGCCAGGGCAAGTGGCCCTGGAGGTGAGACTC 78  
 Db 26 IleValGlyGlnGlnAlaProGlyLysSerLeuArgAla 45  
 QY 79 CACGGCCCATACTGGATGCACTCTGCGGGGCTCCCTACACCCAGTGGGTCTG 138  
 Db 46 AsnGluThrTyrTrpArgHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 65  
 QY 139 ACCCGCGGCGGCTGGTGGAGGCGGCAAGATCTGGCCCTCAGGGTCACTG 198  
 Db 66 ThraAlaAlaIscysValGlyProThrIleAlaAspProAsnLysValArgValGlnLeu 85  
 Qy 199 CGGAGCAGCACCTACTACAGGACCGAGCTGCTGGGTCAAGGATCTGTCAC 258  
 Db 86 ArgLyGlyNttrpArgIleGlySerAspIleLeuAlaValSerArgIleIleThrHis 105  
 QY 259 CACAGGTCTCACGCCAGATGGAGGGACATGCCCTGCTGGAGCTGGAGGAGCG 318  
 Db 105 ProThrPhetyrAlaThrIleInAsnGlyAlaAspIleAlaIleLeuGlnLeuLysAsnPro 125  
 QY 319 GTGAGGCTCTCACCCAGTCACCGTCACACGGTCACTGCCCTGCTCAGAGCTTCCC 378  
 Db 126 ValAsnIleSerSerHisValHisProValSerLeuProProLaserGluThrPhePro 145  
 QY 379 CGGGGATGCCGICCTGGTCACGGCGGGGATGGCAGCATGATGAGGCCCTCCA 438  
 Db 146 SerGlyIleCysTrpValThrGlyLysIleAspAsnValSerLeuIleThrHis 165  
 QY 439 CGGCCATTCTCTGAGGGGAGGTCATAATGAAACACATTGTGACCA 498  
 Db 166 ProProProProLeuLysGluValInvalProValGlnAsnGlnIleCysAspLeu 185  
 QY 499 AAATACCACCTTGCGCTTACACGGAGACGAGTCCCTCCGGAGGCCCTGGTGG 558  
 Db 186 LysTrpIleSerHisGlyValtyrThrGlyAspAsnIleHisIleValArgAspAspMetLeu 205  
 QY 559 TGTGGGGAACACCCGGAGGACTCACCCAGGGGACTCCTGGAGGCCCTGGTGG 618  
 Db 206 CysAlaGlyAsnGlyLysGlyIleSerAspCysGlyGlyAspSerGlyGlyProLeuValCys 225  
 QY 619 AGGGTGAATGGCACCTGGCTGAGGGGGGGGGTGGCTAGCTGGGGAGGGCTGGGCCAG 678  
 Db 226 LysValAsnGlyIleSerLeuGlnAlaGlyIleValSerTrpGlyGluGlyCysAlaLeu 245  
 QY 679 CCCAACGGCTTGCATCTACACCCGGTCACTACTGAGTCACTCAT 738

Db 246 ProAsnArgProGlyIleTyrThrArgValThrIleTyrLeuAspPtpileHisArgTyr 265  
 Qy 739 GTCGCCAA 747  
 Db 266 ValProLys 268

RESULT 15  
 US-08-978-404B-3  
 ; Sequence 3, Application US/08978404B  
 ; Patent No. 5968782  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Stevens, Richard L.  
 ; TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES  
 ; TITLE OF INVENTION: FIBRINOGEN  
 ; NUMBER OF SEQUENCES: 74  
 ; CORRESPONDENCE ADDRESS:  
 ; STREET: 600 Atlantic Avenue  
 ; CITY: Boston  
 ; STATE: MA  
 ; ZIP: 02210-2211  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSQL for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/978,404B  
 ; FILING DATE: 23-Nov-97  
 ; CLASSIFICATION: 435  
 ; PRIORITY APPLICATION DATA:  
 ; APPLICATION NUMBER: 60/032,354  
 ; FILING DATE: 04-DEC-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Plumer, Elizabeth R.  
 ; REGISTRATION NUMBER: 36,637  
 ; REFERENCE/DOCKET NUMBER: B0801/7090  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 617-720-5500  
 ; TELEFAX: 617-720-2441  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 273 amino acids  
 ; STRANDEDNESS: single  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: NO. 5968782e  
 ; US-08-978-404B-3

Alignment Scores:  
 Pred. No.: 7.07e-78 Length: 273  
 Score: 1057.00 Matches: 189  
 Percent Similarity: 82.07% Conservative: 17  
 Best Local Similarity: 75.30% Mismatches: 43  
 Query Match: 72.50% Indels: 2  
 DB: 2 Gaps: 1

US-09-598-982-20 (1-771) x US-08-978-404B-3 (1-273)

Qy 1 GGCCCCCTCGAGAAAAGA-----ATCGGGGGGGTCAGGAGGCCAGGGCAAGTGG 54  
 Db 21 GlyProAlaMetThrArgGluGlyIleValGlyGlnGluLysAsnLysTrp 40

Qy 55 CCCGGCAGGTGGCTGAGAGTCACCGGCCCATACTGGATGACTTCCTGGGGCTCC 114  
 Db 41 ProTrpGlnValSerLeuGlyAlaAsnAspThrTrpMethisPheCysGlyGlySer 60

Qy 115 CTCATGCCACCCCAAGTGCGGCTGAGCCGCGCGTGGGAGCCGAGCTCAAGAT 174  
 Db 61 LeuIleHsProGlnTrpValLeuThrAlaAlaLysCysValGlyProSpValAlaAsp 80

Qy 175 CTGGCCGCCCTAGGTGCAACTCGGGGAGCAGCACCTACTACAGGACCAAGCTG 234  
 Db 81 ProAsnLysValArgValGlnLeuArgLysTrpMethisAspHisLeuMet 100  
 Qy 235 CGGGTCAGGAGATCATCTGACCCACAGTCACACGCCAGATGGAGGGACATC 294  
 Db 101 ThrValSerGlnNileIleHsProAspPheTyrIleValGlnAspGlyAlaAspIle 120  
 Qy 295 GCCCTGCTGAGCTGGAGGTGAGGTGAGGTCAGCCACGTCACGCGTCACCGCTG 354  
 Db 121 AlaLeuLeuLysLeuThrAsnProValAsnIleSerAspTyrValHsProValProLeu 140  
 Qy 355 CCCCTGCCTCAGGACCTCCCGGGATGCCGTCAGTCACUGCGCCGGCAT 414  
 Db 141 ProProAlaSerGluThrProSerGlyIleTrpValAspCysTrpValHsProValProLeu 160  
 Qy 415 GTGGCACATTGAGAGCCTCCACCGCATTCCTCTGAAGGAGGTAAGGCCATA 474  
 Db 161 IleAspAspGlyIlyValAsnLeuProProProHeProLeuIleGluValProLeu 180  
 Qy 475 ATGAAAMACCACTTGTGACGCCAAATACCCACTGGCCGCTACACCGGAAAGACGCTC 534  
 Db 181 IleGluAsnHisLeuCysAspLeuLysTrpHisAspSerCysGlyAspAsnVal 200  
 Qy 535 CGCATCGTCCGTCGACGACATGCTGCGGGACTCATGCCAGGG 594  
 Db 201 HistLeuValArgAspAspMetLeuCysAlaGlyAsnGlyIleThrGlyAspAsnVal 220  
 Qy 595 GACTCCGGAGGCCCTGGTGTCAAGGTGATGGCACCTGGCTGAGGGGGGCTG 654  
 Db 221 AspGlyGlyProLeuValCysLysValGluAspThrPheGlnAlaGlyVal 240  
 Qy 655 AGCTGGGGAGGGCTGRCGCCAGCCCACCGCCCTGGCATCACCCGTCACCTAC 714  
 Db 241 SerTrpGlyGluGlyCysAlaGinProAsnArgProGlyIleTyrThrArgValThrTyr 260  
 Qy 715 TACTGGAGCTGATCCACCACTAATGTCCTCAA 747  
 Db 261 TyrLeuAspPtpileHisHistYrValProLys 271

Search completed: January 31, 2003, 07:00:10  
 Job time : 22.5 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

## SUMMARIES

OM nucleic - protein search, using frame\_plus\_n2p model  
Run on: January 31, 2003, 05:24:20 ; Search time 46 Seconds  
(without alignments)  
4465.793 Million cell updates/sec

Title: US-09-598-982-20  
Perfect score: 1458  
Sequence: 1 gggeccctcgagaaaaaat.....cgtgaagcggccgcgtcgt 771

## Scoring table:

BLOSUM62  
Xgapop 10.0 , Ygapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Egapop 6.0 , Egapext 7.0  
Delop 6.0 , Delext 7.0

## Searched:

908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 1816940

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p model -DEV=x1h  
-Q=/Cgn2\_1/USP1\_spool/US0598982/runat\_27012003\_073715.19505/app\_query.fasta\_1.967  
-DB=A\_Genesed\_101002 -QFMR=fastan -SUFIX=n2p\_rag -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40\_cdi  
-LIST=45 -DOALIGN=200 -THR\_SCORE\_PCT -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTTYPE=pto -NORMEXT=1 -HASPFILE=50000 -MINLEN=0 -MAXLEN=2000000000  
-USER=US0598982@CGN\_1.1.0 @runat\_27012003\_073715\_19505 -NCPU=6 -TCP=3  
-NO\_XLPPY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XCAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A\_Genesed\_101002;\*

1: /SIDS2/gcadata/geneseq/geneseq-emb1/AA1980.DAT:\*

2: /SIDS2/gcadata/geneseq/geneseq-emb1/AA1981.DAT:\*

3: /SIDS2/gcadata/geneseq/geneseq-emb1/AA1982.DAT:\*

4: /SIDS2/gcadata/geneseq/geneseq-emb1/AA1983.DAT:\*

5: /SIDS2/gcadata/geneseq/geneseq-emb1/AA1984.DAT:\*

6: /SIDS2/gcadata/geneseq/geneseq-emb1/AA1985.DAT:\*

7: /SIDS2/gcadata/geneseq/geneseq-emb1/AA1986.DAT:\*

8: /SIDS2/gcadata/geneseq/geneseq-emb1/AA1987.DAT:\*

9: /SIDS2/gcadata/geneseq/geneseq-emb1/AA1988.DAT:\*

10: /SIDS2/gcadata/geneseq/geneseq-emb1/AA1989.DAT:\*

11: /SIDS2/gcadata/geneseq/geneseq-emb1/AA1990.DAT:\*

12: /SIDS2/gcadata/geneseq/geneseq-emb1/AA1991.DAT:\*

13: /SIDS2/gcadata/geneseq/geneseq-emb1/AA1992.DAT:\*

14: /SIDS2/gcadata/geneseq/geneseq-emb1/AA1993.DAT:\*

15: /SIDS2/gcadata/geneseq/geneseq-emb1/AA1994.DAT:\*

16: /SIDS2/gcadata/geneseq/geneseq-emb1/AA1995.DAT:\*

17: /SIDS2/gcadata/geneseq/geneseq-emb1/AA1996.DAT:\*

18: /SIDS2/gcadata/geneseq/geneseq-emb1/AA1997.DAT:\*

19: /SIDS2/gcadata/geneseq/geneseq-emb1/AA1998.DAT:\*

20: /SIDS2/gcadata/geneseq/geneseq-emb1/AA1999.DAT:\*

21: /SIDS2/gcadata/geneseq/geneseq-emb1/AA2000.DAT:\*

22: /SIDS2/gcadata/geneseq/geneseq-emb1/AA2001.DAT:\*

23: /SIDS2/gcadata/geneseq/geneseq-emb1/AA2002.DAT:\*

## DESCRIPTION

Result No.	Score	Query Match	Length	DB ID	Description
1	1333	95.5	249	23 AAU12009	Human beta-II tryp
2	1348	95.2	249	23 AAU12017	Human beta-II tryp
3	1387	95.1	249	21 AAU55011	Human beta-trypas
4	1387	95.1	249	23 AAU12007	Recombinant human
5	1384	94.9	249	23 AAU12011	Human beta-II tryp
6	1384	94.9	249	23 AAU12012	Human beta-II tryp
7	1382	94.8	249	23 AAU12006	Human beta-I tryp
8	1379	94.5	249	23 AAU12010	Human beta-II tryp
9	1379	94.5	249	23 AAU12019	Human beta-II tryp
10	1379	94.5	249	23 AAU12020	Human beta-II tryp
11	1374	94.2	245	23 AAU12013	Human beta-II tryp
12	1374	94.2	249	23 AAU12018	Human beta-II tryp
13	1369	93.9	245	23 AAU12021	Human beta-II tryp
14	1368	93.8	245	23 AAU12008	Human beta-II tryp
15	1368	93.8	274	19 AAW64240	Human beta-II tryp
16	1368	93.8	245	23 AAU12015	Human beta-II tryp
17	1355	93.5	245	23 AAU12016	Human beta-II tryp
18	1365	93.6	245	23 AAU12015	Human beta-II tryp
19	1365	93.5	245	21 AAU55010	Human beta-trypas
20	1353	93.5	245	23 AAU12005	Human mature beta
21	1363	93.5	273	19 AAW64238	Human mast cell tr
22	1363	93.5	273	19 AAW63174	Human mast cell tr
23	1351	93.3	244	20 AAU25925	Human lung mast ce
24	1360	93.3	245	23 AAU12014	Human beta-II tryp
25	1360	93.3	245	23 AAU12023	Human beta-II tryp
26	1350	93.3	245	23 AAU12024	Human beta-II tryp
27	1355	92.9	245	23 AAU12022	Human beta-II tryp
28	1344	92.2	267	19 AAU63175	Human mast cell tr
29	1344	92.2	267	23 AAU84360	Human mast cell tr
30	1344	92.2	267	23 AAU84348	Human protease PR
31	1329	91.2	691	23 AAU12023	Human mast cell tr
32	1274	87.4	275	19 AAU64237	Human mast cell tr
33	1274	87.4	275	19 AAU63173	Human mast cell tr
34	1092	74.9	276	19 AAU64242	Mouse mast cell pr
35	1092	74.9	276	19 AAU63172	Mouse mast cell pr
36	1077	73.9	276	23 ABB05437	Gerbil homologue o
37	1075	73.7	270	19 AAU64239	Murine MCP-7 zymo
38	1057	72.5	273	19 AAU64233	Human protease PR
39	1057	72.5	310	23 AAU12024	Rat homologue of m
40	1055	72.4	274	19 AAU64234	Rat homologue of m
41	1039	71.3	273	19 AAU64235	Pig lung protease
42	1006	69.0	275	21 AAU81826	Amino acid sequenc
43	702	48.1	190	20 AAU68539	Canine homologue o
44	698	47.9	190	20 AAU68540	
45	684.5	46.9	269	19 AAU684236	

## ALIGNMENTS

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.







QY	427	GAGGCCCTCCACCGGATTCCTCTGAGCAGTGAGTCCCCATATGGAAAACCA	486	enzymatically-active, recombinant proteolytic tryptase produced are
Db	141	GluArgLeuProProProProProProLeuLysGlnValProLeuMetGluAsnHis	160	useful as an antigen to generate anti-human tryptase antibodies, and in drug screening for compounds which act as tryptase inhibitors, antagonists, agonists, etc. AAU12009-AAU12024 represent recombinant
QY	487	ATTGTGACGCAAATAACCCACCTGGCCCTAACGGGAGCACGTCGGCATGTCGCT	546	human beta-II tryptase active site mutants.
Db	161	IleCysAspAlaLysTyrIleGlyAlaTyrThrGlyAspAspValArgLeuVal	180	xx
QY	547	GAGCACATGCCTGTCGGGGACACCGGGAGACTCATGCCAGGGCAGTCGGAGGG	606	SQ Sequence 249 AA;
Db	181	AspAspMetLeucysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGly	200	Alignment Scores:
QY	607	CCCTGGTGGCAAGGCGATGGCACCTGGCTGAGGGGGCTGGTCACTGGAGG	666	Pred. No.: 6.86e-97 Length: 249
Db	201	ProLeuValCysLysValAsnGlyThrProLeuGlnAlaGlyValSerTrpGly	220	Score: 1384.00 Matches: 247
QY	667	GCTGTGCCACGCCAACGGGCTGGATCTACACCGCTGTCACCTACTACTTGACTG	726	Percent Similarity: 99.50% Conservative: 1
Db	221	GlycylsAlaGlnProAsnArgProGlyIleTyrrhargValThrTyrIleAspTrp	240	Best Local Similarity: 99.20% Mismatches: 1
Db	241	IleHisHistTyValProLysPro	249	Query Match: 94.92% Indels: 0
Db:	23		Gaps: 0	
RESULT 5				
AAU12011	ID	AAU12011 standard; Protein: 249 AA.		US-09-598-982-20 (1-771) x AAU12011 (1-249)
XX	XX			QY 7 CTCGAGAAGAAGATGTCGGGGTAGGAGCCGCCAGGAGCAAGGGCCCTGGAGGTG 66
AC	AAU12011;			Db 1 LeuGlyLysArgIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnVal 20
XX				QY 67 AGCCGTGAGACTCCAGGCCATACGGAATGCACTTGACTCTGTGCGGGGTCCCAC 126
DT	09-APR-2002 (first entry)			Db 21 SerIleArgValHisGlyIleProTyrrMetIlePheCysGlyGlySerLeuIleHisPro 40
XX				QY 127 CAGTGGTGCAGACGCCGGCGCGGACCGTCAGGTCAGGATCTGGCCGCTC 186
DE				Db 41 GluTrpValIleThrAlaAlaHisCysValGlyProAspValAspLeuAlaLeu 60
XX				QY 187 AGGGGCAACTGCGGAGCACCTCTACTACCGGACCAGCTCTGCGGTGCGCAGG 246
KW				Db 61 ArgValGlnIleArgGluGlnIleHistLeuTyrrGlyAspGlnIleLeuIlePro 80
KW				QY 247 ATCAGTGTGACCCACAGTCTACCCGCCAGTGGACGGACATGCCCTGGAG 306
OS				Db 81 IleIleValIleProGlnIlePheTyrrIleAlaGlnIleGlyIleAlaAspIleAlaLeu 100
Synthetic.				QY 307 CTGGAGGAGCGGTGAGGTCCTCCAGCCACGGTCACCGTCCACGGTCACTGCCCCTCA 366
OS				Db 101 LeuGlyGluProValAlaLysValSerSerHisValIleThrValThrLeuProProAlaSer 120
PN	W020198470-A2.			QY 367 GAGACCTTCCCCCGGGATGCCGTCTGGCTCACGGCCTGGGAGATGACATGAT 426
XX				Db 121 GlutinPheProProGlyMetProCysTrpValIleGlyIleProGlyAspValAspAsp 140
PD	27-DEC-2001.			QY 427 GAGGCCCTCCACCCCATTCCTGAAACGGGAAGGCCATAAAGGAAACCA 486
PF	20-JUN-2001; 2001WO-US19681.			Db 141 GluArgLeuProProProProProLeuLysGlnValProLeuMetGluAsnHis 160
XX				QY 487 ATTGTGACGCAAATAACCCACCTGGCCCTAACGGGAGCACGTCGGCATGTCGCT 546
PR	21-JUN-2000; 2000US-0598982.			Db 161 IleCysAspAlaLysTyrIleGlyAlaTyrThrGlyAspAspValArgLeuAla 180
XX				QY 547 GAGCACATGCCTGTCGGGGACACCCGAGGGACTCATGCCAGGGACTCCGGAGG 606
PA	(PROM-) PROMEGA CORP.			Db 181 AspAspMetLeucysAlaGlyAsnThrArgArgAspSerCysGlnGlyAsp 240
XX				QY 607 CCCCTGGTGGCAAGCTGATGCCCTGGCTGGCGGGCGCTGTCAGCTGGCGAG 666
PI	Maffit M, Niles AL, Haak-Frendscho M;			Db 201 ProLeuValCysLysValAsnGlyIleProLeuGlnAlaGlyValSerTrpGly 220
XX				QY 667 GGCTGTGCCACGCCAACGGCCTGCGCATACACCGCTGTCACCTACTACTGACTGG 726
DR	WPI; 2002-114578/15.			Db 221 GlycylsAlaGlnProAsnArgProGlyIleTyrrIleArgValThrTyrIleAspTrp 240
DR	N-PSDB; AAS2077.			QY 727 ATCAGTGTGACCCACAGTGGACGGGAAGGCCATAAAGGAAACCA 753
XX				Db 241 IleHisHistTyValProLysPro 249
PT				RESULT 6
PT	DNA construct for producing enzymatically-inactive proteolytic			AAU12012
PT	tryptase, comprises DNA sequence encoding proteolytic tryptase having			ID
PT	an active site mutation -			
PS	Claim 8; Page 90-91; 126pp; English.			
CC	The present invention relates to recombinant human proteolytic			
CC	tryptases, active site mutants of these tryptases and the methods for			
CC	producing these. The method involves the production of a DNA expression			
CC	construct comprising a promoter operably linked to a secretion signal			
CC	sequence which is operably linked to a DNA sequence encoding a			
CC	proteolytic tryptase with an active site mutation (the construct drives			
CC	expression of a mature proteolytic tryptase that lacks enzymatic activity			
CC	due to the active site mutation, in hosts transformed to contain the			
CC	construct). The method is useful for producing enzymatically-active			
CC	beta-II tryptase. The active site mutants of proteolytic tryptase provide			
CC	a tool to investigate the structural and functional properties of the			
CC	protease and its enzymatic activity, and for modelling studies. The			





Pred. No. :	1	Length:	249	XX
Score:	1.64e-96	Matches:	247	OS
Percent Similarity:	1378.00	Conservative:	0	Synthetic
Best Local Similarity:	99.20%	Mismatches:	2	XX
Query Match:	99.20%	Indels:	0	WO200198470-A2.
DB:	23	Gaps:	0	PN
US-09-598-982-20 (1-771) x AAU12010 (1-249)				XX
QY	7 CTCGAGAAGAAAGAACGGTCTGGGGCAGAGGCCACAGTGGCAAGTGGCCCTGGCAGTG	Length:	66	XX
Db	LeuGluLysArgIleValGlyProAlaProArgSerIysTrpProtRpglnval	Matches:	20	XX
QY	67 AGCCGTAGACTCCACGCCATACTGGATGCCACTTCTGGGGCTCCCTCATCCCC	Length:	126	XX
Db	LeuGluLysArgIleValGlyProAlaProArgSerIysTrpProtRpglnval	Matches:	20	XX
QY	127 CAGTGGGTCTGACGCCCGGGCTGGGGACGGACGTCAGATCTGGCCCTGGCAGTG	Length:	186	XX
Db	LeuGluLysArgIleValGlyProAlaHisCysValGlyProAspValLysAspLeuAlaLeu	Matches:	60	XX
QY	41 GlnTrpValLeuThrAlaAlaHisCysValGlyProAspValLysAspLeuAlaLeu	Length:	60	XX
Db	LeuGluLysArgIleValGlyProAlaProArgSerIysTrpProtRpglnval	Matches:	60	XX
QY	187 AGCTGCAACTGGGACAGCACCTCTACTACAGGACAGCTGGGGTCAGATCTGGCCCTGGCAGTG	Length:	246	XX
Db	LeuGluLysArgIleValGlyProAlaProArgSerIysTrpProtRpglnval	Matches:	60	XX
QY	61 ArgValGlnLeuArgGluGlnHisLeuTyryGlnAspGlnLeuLeuProValSerArg	Length:	80	XX
Db	LeuGluLysArgIleValGlyProAlaProArgSerIysTrpProtRpglnval	Matches:	80	XX
QY	61 ArgValGlnLeuArgGluGlnHisLeuTyryGlnAspGlnLeuLeuProValSerArg	Length:	80	XX
Db	LeuGluLysArgIleValGlyProAlaProArgSerIysTrpProtRpglnval	Matches:	80	XX
QY	247 ATCATGTCGACCCACAGTCTACGCCAGATGGAGGAGCATCCCTCATCCCC	Length:	306	XX
Db	LeuGluLysArgIleValGlyProAlaProArgSerIysTrpProtRpglnval	Matches:	100	XX
QY	307 CTGGAGGAGCGGGAGGGCTCCAGCACCGTCAACCGTCAACCTGGCCCTGGCTCA	Length:	366	XX
Db	LeuGluLysArgIleValGlyProAlaProArgSerIysTrpProtRpglnval	Matches:	120	XX
QY	101 LeuGluLysProValLysLysSerHisValIleValThrAlaAlaLeuLeuGlu	Length:	426	XX
Db	LeuGluLysArgIleValGlyProAlaProArgSerIysTrpProtRpglnval	Matches:	120	XX
QY	367 GAGACTTCCCCGGGGGCGCGCTGGTCTGGCGGGCATGGACATGATGATG	Length:	426	XX
Db	LeuGluLysArgIleValGlyProAlaProArgSerIysTrpProtRpglnval	Matches:	120	XX
QY	121 GluThrPheProProGlyMetProCysIleValThrGlyTrpGlyAspValAspAsp	Length:	140	XX
Db	LeuGluLysArgIleValGlyProAlaProArgSerIysTrpProtRpglnval	Matches:	120	XX
QY	427 GAGCCCTCCACGCCATTCCTGAAAGCAGTGAAAGTCCCCATAATGAAAACCAC	Length:	486	XX
Db	LeuGluLysArgIleValGlyProAlaProArgSerIysTrpProtRpglnval	Matches:	160	XX
QY	141 GluArgLeuProProProheProLeuIysGlnVallysValProIleMetGluAsnHis	Length:	160	XX
Db	LeuGluLysArgIleValGlyProAlaProArgSerIysTrpProtRpglnval	Matches:	160	XX
QY	487 ATTGTGACCAAATTACCACTGGGCCACGGGAGGCTGGCCTGGCAGCTGGCCT	Length:	546	XX
Db	LeuGluLysArgIleValGlyProAlaProArgSerIysTrpProtRpglnval	Matches:	180	XX
QY	161 IleCysAspAlaLysTyrIleLeuIleGluIleTyrGlyAspAspValArgIleValArg	Length:	180	XX
Db	LeuGluLysArgIleValGlyProAlaProArgSerIysTrpProtRpglnval	Matches:	180	XX
QY	547 GACGCACATGCTGTGGCCGACACCCGGAGGAACGGGACTCATGCCAGGGGG	Length:	606	XX
Db	LeuGluLysArgIleValGlyProAlaProArgSerIysTrpProtRpglnval	Matches:	200	XX
QY	181 AspAspMetIeuCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGly	Length:	200	XX
Db	LeuGluLysArgIleValGlyProAlaProArgSerIysTrpProtRpglnval	Matches:	200	XX
QY	607 CCCCTGGTGTGCAAGCTGGAACTGGCCTGGCTGGAGGGCGGCGTCACTGGCGAG	Length:	666	XX
Db	LeuGluLysArgIleValGlyProAlaProArgSerIysTrpProtRpglnval	Matches:	220	XX
QY	201 ProLeuValCysLysValAsnGlyThrIlePheLeuGlnAlaGlyValValSerIysTrpGlyGlu	Length:	220	XX
Db	LeuGluLysArgIleValGlyProAlaProArgSerIysTrpProtRpglnval	Matches:	220	XX
QY	667 GGCGTGTGCCAGCCAACGGCTGGCACATCACCCGGTCACTACTGGACTG	Length:	726	XX
Db	LeuGluLysArgIleValGlyProAlaProArgSerIysTrpProtRpglnval	Matches:	240	XX
QY	221 GlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrp	Length:	753	XX
Db	LeuGluLysArgIleValGlyProAlaProArgSerIysTrpProtRpglnval	Matches:	240	XX
Db	241 IleHisHistidylValProLysPhePro	Length:	249	XX
RESULT 9				XX
AAU12019				XX
ID - AAU12019 standard; Protein; 249 AA.				XX
XX AAU12019;				XX
XX DT 09-APR-2002 (first entry)				XX
DE Human beta-II tryptase active site mutant S194A #5.				XX
KW Human; proteolytic tryptase; protease; recombinant beta-II tryptase; enzyme; mutant; mutant.				XX

QY	CTGGAGCAGCGGTGAAGGTCRCCACGCCAACGGCACCCCTGGCTCA	366	CC
Db	IeuglugluProvalAsnValserSerHisValHisThrLeuProProAlaser	120	CC
QY	367 GAGACCTCCCCGGGAGACCGGGTACTGGCTGGGAGATGGACATGAT	426	CC
Db	121 GluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspAsp	140	CC
QY	427 GAGCGGCCTCCACCGCATTCCTCAAGCAGGTGAGGTCGCCATATGGAAACAC	486	CC
Db	141 GluArgLeuProProProPheProLeuLysGlnValProleMetGluAsnHis	160	CC
QY	487 ATTGCTGAGCAAATTACCACTTGCGCTACACGGGAGACGCCATCGCGCT	546	CC
Db	161 IlecyAspAlaLysTyrHistLeuGlyAlaTyrThrGlyAspAspValArg	180	CC
QY	547 GACGACATGCTGTCGGGGACACCGGAGGGACTCATGCCAGGGGACTCGGGAGG	606	CC
Db	181 AspAspMetLeuCysAlaGlyAsnThrArgGlyAspSerCysGlnGlyAspAlaGly	200	CC
QY	607 CCCCTGGTGTGAGGGAATGCACTTGCGTCAGCGGGGGTGGCAGCAGGGAG	666	CC
Db	201 ProLeuValCysLysValAsnGlyThrPheLeuGlnAlaGlyValValSerTrpGlyGlu	220	CC
QY	667 GGCTGCCCCAGCCCACGGCCTGGCATCACCCGGTGTCACTACTTGACTG	726	CC
Db	221 GlyCysAlaGlnInProAsnArgProGlyIleTyrThrArgValThrTyrLeuAspTrp	240	CC
QY	727 ATCCACCACTAATGCCCAAAAGCCG	753	CC
Db	241 IleHisHistYryValProLysLysPro	249	CC
RESULT 10			Sequence 249 AA;
AAU12020	AAU12020 standard; Protein: 249 AA.		
DT	AAU12020		
XX	AC		
XX	AC		
XX	AAU12020;		
XX	09-APR-2002 (first entry)		
XX	Human beta-II tryptase active site mutant S194A #6.		
KW	Human; proteolytic tryptase; protease; recombinant beta-II tryptase;		
KW	enzyme; mutant; mutein.		
XX	Homo sapiens.		
OS	Synthetic.		
XX	WO200198470-A2.		
XX	21-JUN-2000; 2000US-0598982.		
PD	27-DEC-2001.		
XX	PF 20-JUN-2001; 2001WO-US19681.		
PR	21-JUN-2000; 2000US-0598982.		
PA	(PROM-) PROMEGA CORP.		
XX	PI Maffit, M., Niles AL, Haak-Frendscho M;		
XX	DR WPI; 2002-114578/15.		
XX	N-PSDB; AAS20786.		
DNA construct for producing enzymatically-inactive proteolytic tryptase, comprises DNA sequence encoding proteolytic tryptase, having an active site mutation -			
PS Claim 8; Page 114-115; 126pp; English.			
The present invention relates to recombinant human proteolytic tryptases, active site mutants of these tryptases and the methods for			
CC	producing these. The method involves the production of a DNA expression construct comprising a promoter operably linked to a secretion signal sequence which is operably linked to a DNA sequence encoding a proteolytic tryptase with an active site mutation (the construct drives expression of a mature proteolytic tryptase that lacks enzymatic activity due to the active site mutation), in hosts transformed to contain the construct. The method is useful for producing enzymatically-active beta-II tryptase. The active site mutants of proteolytic tryptase provide a tool to investigate the structural and functional properties of the protease and its enzymatic activity, and for modelling studies. The enzymatically-active, recombinant proteolytic tryptase produced are useful as an antigen to generate anti-human tryptase antibodies, and in drug screening for compounds which act as tryptase inhibitors, antagonists, agonists, etc. AAU12009-AAU12024 represent recombinant human beta-II tryptase active site mutants.		
US-09-598-982-20 (1-771) x AAU12020 (1-249)			
QY	7 CTCGAGAAAGAATGTCGGCTTAGGGAGCCACGGGCAAGTGGCCTGGAGGTG	66	CC
Db	1 IeuglugluArgIleValGlyIleGlyTrpGlyGlnAlaProTrpSerIleLeuHisPro	20	CC
QY	67 AGCCCTGAGCTCCAGGCCATACTGGATGCCACTCTGGGGGCTTCATCCACCC	126	CC
Db	21 SerIeuArgValHisGlyProTyrTrpMetHisPheCysLyGlySerIleLeuHisPro	40	CC
QY	127 CAGTGGTGTGACGCCGGGGTGGCTGGACCGACCTCANGATCTGGCCCTCTC	186	CC
Db	41 GluIleValLeuThrAlaAlaHisCysValGlyProAspValLysAspLeuAlaLeu	60	CC
QY	187 AGGGGCAACTGCGGAGCAGCACCTCTACTACCGGACAGCTCTGGGGTCGACAG	246	CC
Db	61 ArgGlyAlaIleValGlyGluGlyIleValAspIleLeuLeuProValSerArg	80	CC
QY	247 ATCATCTGGACCCACAGTCTACACGCCAGATGGACGGAGATGCCCTGGGG	306	CC
Db	81 IleIleValHisProGlnInPheTyrThrAlaGlnIleGlyValAspIleLeuLeuGlu	100	CC
QY	307 CTGGGGGAAACGGGTAAAGGCTCCGCCACGCTCCACCTCACCTGCCCTGGCTCA	366	CC
Db	101 LeuglugluProvalAsnValSerSerHisValHisThrValThrLeuProProAlaser	120	CC
QY	367 GAGACCTCCCCGGGATGCCGGCTGGTACTGGCTGGGGATGAGCATGAT	426	CC
Db	121 GluArgLeuProProGlyMetProCysTrpValThrGlyIlePrglyAspAsp	140	CC
QY	487 ATTGCTGAGCAAATTACCACTTGCGCTACAGGGAGACGAGTCGCCATCTCCGT	546	CC
Db	161 IlecyAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgLys	180	CC
QY	547 GACGACATGCTGTCGGGGACACCCGGAGGACTCATGCCAGGGACTCCGGAGG	606	CC
Db	181 AspAspMetLeuCysAlaGlyAsnThrArgGlySerCysGlnGlyIleValGly	200	CC
QY	607 CCCCTGGTGTGCAAGGTGAATGGCACCTGCTGAGCGGGGGTCACTGGGGAG	666	CC
Db	201 ProLeuValCysLysValAsnGlyThrPheLeuGlnAlaGlyValValSerTrpGlyGlu	220	CC
Ov	667 GGCTGCCCCAGGCCAACGGCTGGCATCACCCGGTACTACTTGACTGG	726	CC



PF 20-JUN-2001; 2001WO-US19681.  
 XX  
 PR 21-JUN-2000; 2000US-0598982.  
 XX  
 PA (PROM-) PROMEGA CORP.  
 XX  
 PI Maffit M, Niles AL, Haak-Frendscho M;  
 XX DR WPI; 2002-114578/15.  
 XX N-FSDB; AAS20784.  
 PT DNA construct for producing enzymatically-inactive proteolytic  
 trypsin, comprises DNA sequence encoding proteolytic trypsin having  
 an active site mutation -  
 PS Claim 8; Page 109-110; 126pp; English.  
 CC  
 CC The present invention relates to recombinant human proteolytic  
 trypsinases, active site mutants of these trypsinases and the methods for  
 producing these. The method involves the production of a DNA expression  
 construct comprising a promoter operably linked to a secretion signal  
 sequence which is operably linked to a DNA sequence encoding a  
 proteolytic trypsinase with an active site mutation (the construct drives  
 expression of a mature proteolytic trypsinase that lacks enzymatic activity  
 due to the active site mutation, in hosts transformed to contain the  
 construct). The method is useful for producing enzymatically-active  
 beta-II trypsin. The active site mutants of proteolytic trypsin provide  
 a tool to investigate the structural and functional properties of the  
 protease and its enzymatic activity, and for modeling studies. The  
 enzymatically-active, recombinant proteolytic trypsin produced are  
 useful as an antigen to generate anti-human trypsin antibodies  
 and in drug screening for compounds which act as trypsin inhibitors,  
 antagonists, agonists, etc. AAU12009-AAU12024 represent recombinant  
 CC human beta-II trypsin active site mutants.  
 CC  
 SQ Sequence 249 AA;  
 Alignment Scores:  
 Pred. No.: 3.93e-96 Length: 249  
 Score: 1374.00 Matches: 246  
 Percent Similarity: 98.80% Conservative: 0  
 Best Local Similarity: 98.80% Mismatches: 3  
 Query Match: 94.24% Indels: 0  
 DB: 23 Gaps: 0  
 US-09-598-982-20 (1-771) x AAU12018 (1-249)  
 QY 7 CTCGGAAAGAATGCTCGGGGCTAGGAGCCCCAGAGCAAGTGGCCCTGGCAGGTG 66  
 DB 1 LeuGluLysArgIleValGlyIleGlnGluAlaProArgSerIleTrpProGlnVal 20  
 QY 67 AGCCCTGAGACTCCAGGCCCATCTGGACTCTGGGCTCTCATCCACCC 126  
 DB 21 SerIleArgValHisGlyProTyrTrpMetHisPheCysGlyGlySerIleIleHisPro 40  
 QY 127 CAGTGGGTGCTGACGGCGGGCTGGGACGGACGTCAGAGATCTGGCCCTC 186  
 DB 41 GluIleValLeuThrAlaAlaHisCysValGlyProAspLeuAspIleAlaLeu 60  
 QY 187 AGGGTGCACACTGCGGAGCACCTCTACTACCAAGGACAGCTGCTGGGTCAGCAG 246  
 DB 61 ArgValGlnIleLeuArgGluGlnHisLeuTyroGlnAspGlnLeuLeuProValSerArg 80  
 QY 247 ATCATGGTGACCCGAGTCTACCCGCCAGATGGACGACATGCCCTGGAG 306  
 DB 81 IleLeuValHisProGlnPheTyroThraAlaGlnIleGlyAlaAlaIleLeuLeuGlu 100  
 QY 307 CTGGAGGAGCCGCTGAGGTCTCCAGCCACGGTCACCCCTGCCCTGGCTCA 366  
 DB 101 LeuGluGluProValAsnValSerSerHisValIleThrValIleLeuProProLys 120  
 QY 367 GAGACCTCCCGGGGAGGCCGCTGGTCACTGGCTGGGGGAGCTGGGAGATGAT 426  
 CC  
 DB 121 GluThrPheProProGlyMetProCystPvalThrGlyTrpGlyAspValAspAsp 140  
 QY 427 GAGGCCCTCCACGCCATTCTCTGAAGCAGGTCCCATAATGGAAACAC 486  
 DB 141 GluArgLeuProProProPheProLeuLysGlnValProIleMetGluAsnHis 160  
 QY 487 ATTTGGTACCCAAATACCACCTGGCCTACAGGGAGACGAGCTGGCCATGGCTCT 546  
 DB 161 IleGlyAspAlaLysTyrIleGlyIleGlyAspAspValArgLeuValIle 180  
 QY 547 GAGGACATGCTGTCGCCGGAACACCCGAGGAGCTCATGCCAGGGCACTCGGAGG 606  
 DB 181 AspAspMetLeucYsAlaGlyAsnThrArgGlyAspSerCysGlyAspSerIleGly 200  
 QY 607 CCCCGGTCGACGCTGATGGACTGGTGCAGGGGGGGTCAAGCTGGGGAG 666  
 DB 201 ProLeuValCysLysValAsnGlyThrIleLeuGlnAlaGlyIleValSerTrpGlyIu 220  
 QY 667 GGCTGTGCCAGCCCACCCGCTGGCATCTACACCGGTACACTACTGGACTGG 726  
 DB 221 GLYCYSAlaGlnProAsnArgProGlyIleTyrIleArgValThrTyrLeuAspTrp 240  
 QY 727 ATCCACCACTATGTCGCCAAAGCCG 753  
 DB 241 IleHistIleTyrValProLysLysPro 249  
 RESULT 13  
 AAU12021  
 ID AAU12021 standard; Protein: 245 AA.  
 XX  
 AC AAU12021;  
 XX DT 09-APR-2002 (first entry)  
 XX DE Human beta-II trypsin active site mutant H44A #4.  
 XX KW Human; proteolytic trypsin; protease; recombinant beta-II trypsin;  
 XX enzyme; mutant; mutein.  
 OS Homo sapiens.  
 OS Synthetic.  
 XX PN WO200198470-A2.  
 XX PD 27-DEC-2001.  
 XX PP 20-JUN-2001; 2001WO-US19681.  
 XX PR 21-JUN-2000; 2000US-0598982.  
 XX PA (PROM-) PROMEGA CORP.  
 XX PI Maffit M, Niles AL, Haak-Frendscho M;  
 XX DR WPI; 2002-114578/15.  
 XX N-FSDB; AAS20784.  
 PT DNA construct for producing enzymatically-inactive proteolytic trypsinase, comprises DNA sequence encoding proteolytic trypsin having an active site mutation -  
 PT an active site mutation -  
 PS Claim 40; Page 117-118; 126pp; English.  
 CC  
 CC The present invention relates to recombinant human proteolytic trypsinases, active site mutants of these trypsinases and the methods for producing these. The method involves the production of a DNA expression construct comprising a promoter operably linked to a secretion signal sequence which is operably linked to a DNA sequence encoding a proteolytic trypsinase that lacks enzymatic activity expression of a mature proteolytic trypsin that lacks enzymatic activity due to the active site mutation, in hosts transformed to contain the construct). The method is useful for producing enzymatically-active beta-II trypsin. The active site mutants of proteolytic trypsin provide



QY 139 ACCGGCCGCGGTGGACCGGACTGAGGAGCTGGCCCTCAGGTGCAACTG 198  
 ||||||| ||||| ||||| ||||| ||||| |||||  
 Db 41 ThralAlaAlaHisCysValGlyProAspValLeuAlaLeuArgValGlnLeu 60  
 QY 199 CGGGACGACGACTCTACTACAGGACGAGCTGCGCCGTCAGAGGTCACTG 258  
 ||||||| ||||| ||||| ||||| ||||| |||||  
 61 ArgGluGlnHisLeuTyrrTyrgInaspGinLeuLeuproValSerArgIleLeuValHis 80  
 Qy 259 CCACGTTCAACACGCCAGATGGAGGACATGCCCTGCGAGCTGGAGGCC 318  
 ||||||| ||||| ||||| ||||| ||||| |||||  
 Db 81 ProGlnPhenPheTyrrThraLagInileYaaSpIeAlaLeuLeuGlueGluGluPro 100  
 ||||||| ||||| ||||| ||||| ||||| |||||  
 QY 319 GTGAAGGTCTCCAGGCACGTCACAGGGTACCCCTGCCCTGCTCAGAACCT 378  
 ||||||| ||||| ||||| ||||| ||||| |||||  
 121 ProGlyMetProCysTrpValIthrDlyTrpGlyAspValAspAspGluArgLeuPro 140  
 Db 101 ValIysValSerSerHisValHisthValThrLeuProProAlaSerGluThrPhePro 120  
 QY 379 CCGGGATGCGTGTGGGCACTGCTGSGGCGTGTSGACAATGATGAGCTGGACGCA 438  
 ||||||| ||||| ||||| ||||| ||||| |||||  
 122 ProGlyMetProCysTrpValIthrDlyTrpGlyAspValAspAspGluArgLeuPro 140  
 Db 141 ProProHpheProLeuLysGlnValIysValProIemMetIluAsnHisIleCys 160  
 ||||||| ||||| ||||| ||||| ||||| |||||  
 QY 499 AAATACACCTGGCCCTACACGGAGAGCAGCTGGCATCGTGTGACAGATGCTG 558  
 ||||||| ||||| ||||| ||||| ||||| |||||  
 161 LysTrpHistGluGlyAlaTyrrThreGlyAspPvaLargIleValArgAspMetLeu 180  
 Db 439 CCGCATTCCTCAGAACGGACTCATGGAAACACATTTGGACGCGA 498  
 ||||||| ||||| ||||| ||||| ||||| |||||  
 181 CysAlaGlyAsnThrArgGaspSerCysGlnGlyAspSerGlyGlyProLeuValCys 200  
 QY 559 TGTGGGGGACACCCGGAGGACTCATGGCAGGCGACTCCGGGGCCCTGTGTC 618  
 ||||||| ||||| ||||| ||||| ||||| |||||  
 182 CysAlaGlyAsnThrArgGaspSerCysGlnGlyAspSerGlyGlyProLeuValCys 200  
 Db 619 AAGGGATAATGCACTGGCTGACGGGGGGTGAGCTGGGGGGCTGCCCCAG 678  
 ||||||| ||||| ||||| ||||| ||||| |||||  
 201 LysAlaAsnGlyThrIlePheLeuGlnAlaGlyValSerIrpGlyGluGlyValGln 220  
 Db 679 CCCANCCGCCCTGGATCTACACCGGTGCACTACTGGAGTCCACCACTAR 738  
 ||||||| ||||| ||||| ||||| ||||| |||||  
 Db 221 ProAsnArgProGlyIleTyrrThrArgValThrTyrrTyreLeuAspTrpIleHistYr 240  
 QY 739 GTCCCCAAAAGCCC 753  
 ||||||| ||||| |||||  
 Db 241 ValProLySPro 245  
 RESULT 15  
 AAW64240  
 ID AAW64240 standard; protein; 274 AA.  
 XX  
 XX AAW64240;  
 XX  
 DT 24-NOV-1998 (first entry)  
 XX  
 DE Human mast cell tryptase II/beta.  
 XX  
 KW Mast cell tryptase II/beta; human; MCP-7; mast cell protease 7;  
 KW blood clot; anticoagulant; myocardial infarction; reocclusion;  
 KW thromboembolism; cerebral embolism; thrombosis; therapy.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..18 /label= Sig\_peptide  
 FT Peptide 19..28 /label= pro\_peptide  
 FT Protein 29..274 /label= Mat\_protein  
 XX  
 PN W09824886-A1.  
 PD 11-JUN-1998.

PS 25-NOV-1997; 97WO-US21620.  
 PR XX 04-DEC-1996; 96US-0032354.  
 PA (BGHM ) BRIGHAM & WOMENS HOSPITAL.  
 PI XX Stevens RL;  
 DR XX WPI: 1998-333308/29.  
 N-PSDB; AAV4330.

PT New compositions containing trypase-7, e.g., mouse mast cell protease-7, are used to treat clot formation in e.g., myocardial infarction, reocclusion following angioplasty or pulmonary thromboembolism.

PS Disclosure: Page 66-67; 92pp; English.

XX This is the deduced amino acid sequence of human mast cell trypase II/beta (see also AAV4330). The invention provides: compositions comprising an isolated trypase-7 that may include chimeric proteins that contain (a) a human trypase for all but the active site region and (b) the substrate-binding pocket of mouse trypase-7 or its homologues (see AAW6423-39); a method for treating a blood clot by administering a nucleic acid molecule that codes for a trypase-7, or an expression product, to decrease fibrinogen activity; a nucleic acid encoding a serine protease (SP); and a method of producing a mature SP by expressing the inactive zymogen in a host cell, and cleaving the enterokinase susceptibility domain. The trypase-7 polypeptides can be used to treat disorders mediated by undesirable thrombus clot formation such as myocardial infarction and reocclusion following angioplasty of blood clots associated with pulmonary thromboembolism, deep vein thrombosis, cerebral embolism, renal vein and peripheral arterial thrombosis. They are also useful for all surgical procedures that require decreased blood clots.

XX SQ Sequence 274 AA;

Alignment Scores:  
 Pred. No.: 1.13e-95 Length: 274  
 Score: 1368.00 Matches: 244  
 Percent Similarity: 99.59% Conservative: 0  
 Best Local Similarity: 99.59% Mismatches: 1  
 Query Match: 93.83% Indels: 0  
 DB: 19 Gaps: 0

US-09-598-982-20 (1-771) x AAW64240 (1-274)

QY 19 ATCGTGGGGTCAGGAGCCCCAGAGAAGTGGCCCTGGCAGGTGACCTGAGACT 78  
 ||||||| ||||| ||||| ||||| |||||  
 Db 30 IleValGlyGlyGlnAlaProArgSerIleTrpProArgSerIleValSerIleArgVal 49  
 QY 79 CACGGCCCATCTGGTGGACTCTCGGGGCTCCCTACCCACCCCAAGGGCTG 138  
 ||||||| ||||| ||||| ||||| ||||| |||||  
 Db 50 HisGlyProIleTrpMetHisPheCysGlyGlySerIleIleHisProGlnTrpValLeu 69  
 QY 139 ACCGGCGCGGGTGGCGACCGGAGCTGAGGATCTGGCCGCTCAGGTGCAACTG 198  
 ||||||| ||||| ||||| ||||| ||||| |||||  
 Db 70 ThralAlaAlaHisCysValGlyProSpValLeuAlaLeuArgValGlnLeu 89  
 QY 199 CGGGAGCACCTACTACAGGACGCTGGCCGGTCAGGAGTCATCGTGCAC 258  
 ||||||| ||||| ||||| ||||| |||||  
 Db 90 ArgGluGlnHisLeuTyrrTyrgInaspGlnLeuLeuproValSerArgIleLeuValHis 109  
 QY 259 CCACAGTCTCACCCAGATCGAGCAGATCGCAGCTGGCCCTGCTGAGCAGGCC 318  
 ||||||| ||||| ||||| ||||| |||||  
 Db 110 ProGlnPhenPheTyrrThraLagInileYaaSpIeAlaLeuLeuGlueGluGluPro 129  
 QY 319 GTGAAGGTCTCAGGCACGTCACAGGGCACCCGCCCCCTGGCTCACAGACCT 378  
 ||||||| ||||| ||||| ||||| |||||  
 Db 130 ValIysValSerSerHisValHisthValThrLeuProProAlaSerGluThrPhePro 149

OY	379	CCGGGGATGCCGCTGGCTGGGGCATGTGGCGCCTGCCA	438
Db	150	ProGlyMetProGlySystPvalInhrGlyTrpGlyAspValAspAspGluArgLeuPro	169
QY	439	CCGCATATTCTGAGCAGGCTGCCATAATGAAACACATTTGTGAGCA	498
Db	170	ProProProProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla	189
QY	499	AATACCACCTGGCCTACGGGAGAACGTCGCTGGCTGAGCACAGCTG	558
Db	190	LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetIeu	209
QY	559	TGTCCCCGGAACACCCGGGGACTCTGGCAGGGCAGCTCGGAGGCCCTGGITGC	618
Db	210	CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyIleProLeuCys	229
QY	619	AAGGTGAAAGGCACTGGCTGAGCGGGCGTGGCTGAGCTGGCGAGGCGTGGCCAG	678
Db	230	LysValAsnGlyThrIlePheGlnAlaGlyIleValSerIleGlyIleGlyCysAlaGin	249
QY	679	CCAAACGGCCGCACTACACCCGTCACCTACTACTTGACTGATGCCACCAT	738
Db	250	ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspIleHistYr	269
QY	739	GTCCTCAAAAGCCG	753
Db	270	ValProLysLysPro	274

Search completed: January 31, 2003, 06:53:08  
Job time : 59 secs

GenCore version 5.1.3  
copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model  
Run on : January 31, 2003, 06:54:15 : Search time 31 Seconds  
(without alignments)  
1003.720 Million cell updates/sec

Title: US-09-598-982-20  
Perfect score: 1458

Sequence: 1 gggccccctcgagaaaaaaat.....cgtgaacggccgcgtcg 771

Scoring table: BLOSUM62  
Xgapop 10.0 , Ygapext 0.5  
Ygapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delect 7.0

Searched: 122226 seqs, 20178351 residues

Total number of hits satisfying chosen parameters: 244452  
Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%, Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEl frame+\_n2p.model -DEV=x1h  
-DB=Published Applications AA -OFMM=fastan -SUFFIX=n2p\_rapp -MINMATCH=0.1  
-Q=/sgn2\_/\_USPO\_spool/US09598982/runat\_27012003\_073718\_19634/app\_query.fasta\_1 967

-LOOPCL=0 -LOOPEXT=0 -UNITS=bts -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40\_cdi -LIST=45 -DOCALIGN=200 -TIR\_SCORE=prt -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=LOCAL -OUTFILE=pt0 -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=20000000 -USER=US09598982@GN\_1\_1\_3@runat\_27012003\_073718\_19634  
-NCPU=6 -ICPU=3 -NO\_XLPPY -NO\_MMAPP -LARGEQUERY -NRC\_SCORES=0 -WAIT -LONGLOG  
-DEV TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YAPEXT=0.5 -DELOP=6 -DELECT=7

Database : Published\_Applications\_AA: \*

```

1: /cgn2_6/ptodata/2/pubpaas/US08_NEW.PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaas/PCT_NEW.PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaas/US06_NEW.PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaas/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaas/US07_NEW.PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaas/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaas/PTOS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaas/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaas/US09_NEW.PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaas/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaas/US10_NEW.PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaas/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaas/US60_NEW.PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaas/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### ALIGNMENTS

```

RESULT 1
US-09-900-754-2
; Sequence 2, Application US/09900754
; Patient No. US/0020026654AI
; GENERAL INFORMATION:
; APPLICANT: Allen, Michael W.
; APPLICANT: Leviten, Michael W.
; TITLE OF INVENTION: TRANSGENIC MICE CONTAINING TRYPTASE GENE
; TITLE OF INVENTION: DISRUPTIONS
; FILE REFERENCE: R-372
; CURRENT APPLICATION NUMBER: US/09/900,754
; CURRENT FILING DATE: 2001-07-05
; PRIORITY APPLICATION NUMBER: US 60/216,109
; PRIORITY FILING DATE: 2000-07-06
; PRIORITY APPLICATION NUMBER: US 60/223,172
; PRIORITY FILING DATE: 2000-08-07
; PRIORITY APPLICATION NUMBER: US 60/244,111
; PRIORITY FILING DATE: 2000-10-26
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 311
; TYPE: PTM
; ORGANISM: Mus musculus
US-09-900-754-2
Alignment Scores:
Pred. No.: 5.56e-33 Length: 311
```



RESULT 3  
US-10-041-006A-7  
; Sequence 7, Application US/10041006A  
; Patent No. US2002168754A1  
; GENERAL INFORMATION:  
; APPLICANT: Andrade-Gordon, Patricia  
; APPLICANT: Barrow, Andrew  
; APPLICANT: QI, Jian-shen  
; TITLE OF INVENTION: DNA encoding the novel human serine  
; TITLE OF INVENTION: protease T  
; FILE REFERENCE: ORT-1032  
; CURRENT APPLICATION NUMBER: US/10/041,006A  
; CURRENT FILING DATE: 2002-01-07  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO: 7  
; LENGTH: 290  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-10-041-006A-7

Alignment Scores:  
Pred. No.: 4.89e-30 Length: 290  
Score: 563.50 Matches: 110  
Percent Similarity: 61.85% Conservative: 44  
Best Local Similarity: 44.18% Mismatches: 86  
Query Match: 38.65% Index: 9  
DB: 9 Gaps: 4

US-09-598-982-20 (1-771) x US-10-041-006A-7 (1-290)

QY 16 AGAAATGTCGGGGCAGAGGAGCCAGGAGCAAGTGCCCTGGCAGGTGAGCTGAG 75  
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
Db 34 ArgMetValGlyGlyGlnAspThrGlnGluGlyGutrpProGlnValSerIleGin 53

QY 76 GTCCATCGGCCATACTGGATGCACTCTCGGGGCTCCATCACCCCCAGTGGGG 135  
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
Db 54 ArgAsnGly-----SerHisPheCysGlyGlySerIleAlaGluInTrpVal 70

QY 136 CTGACGCCGGGGCTGGCTGGACGGAGCTAACGATCTGGCCCTCAGGGCAA 195  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
Db 71 LeuThrAlaAlaHisCysPhe--ArgAsnThrSerGluThrSerLeuVyrGlnValLeu 89

QY 196 CTGCGGGAGGAGCACCTCTACTAC-----CAGGACCAGCTCTGCGGGTAGCAG 246  
|||:|||||:|||||:|||||:  
Db 90 LeuGlyAlaArgGlnLeuValGlnProGlyProHisAlaMetTyRalaArgValArgGin 109

QY 247 ATCAATGTCGACCCACAGTTCTACACCCGCAGATGGACGGCATGCCCTCTGGAG 305  
:::|||:|||||:|||||:|||||:|||||:|||||:  
Db 110 ValGluSerAsnProLeuSerAsnProLeuSerAsnProLeuSerAsnProLeuValGlu 129

QY 307 CTGGAGGAGCCGCTGAGGCTCCACGCCAGCTCCACCGTCACCTGCCCTGCCCTCA 366  
|||||:|||||:|||||:|||||:  
Db 130 LeuGluAlaProValProPheThrAsnTyPheLeuProValCysLeuProAspPProSer 149

QY 367 GAGACACTCCCCGGGATGCCGIGCTGGTCAGTGGCTGGGGGATGTTGAGCATGAT 426  
|||:|||||:  
Db 150 ValIlePheGluLysMetAsnCysTrpValThrGlyTrpGlySerProSerGluGlu 169

QY 427 GAGCCCTCCACCCCATTCCTCTGAGCAGGGGAAGTCCCATAATGAAAACAC 486  
|||:|||||:|||||:|||||:|||||:  
Db 170 AspLeuLeuProGluProArgIleGluGlnLysIleAlaValProIleAspThrPro 189

QY 487 ATTGTGACCCAATTACACCTTGCGCTACAGGGAGACGAGGCCATCCCGT 546  
|||:|||||:  
Db 190 LysCysAsnIleLeuLeuTyrsLysPhePheGlyIleGlnProLySthrIleLys 209

QY 547 GACGAGATGCTGTGCGCCGG-----AACACCCGAGGACTAAGCCAGGGGACTCC 600  
:::|||||:|||||:  
Db 210 AsnAspMetIleCysAlaGlyPheGluGluGlyLysLysAspAlaCysLysGlyAspSer 229

QY 601 GGAGGCCCCCTGGTGTGCAAGGTGATGGCACCTGCTGCAGGGGGTGGTCAGCTGG 660

RESULT 4  
US-10-028-072-222  
; Sequence 222, Application US/10028072  
; Publication No. US20030004311A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Destroyer, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Wang  
; TITLE OF INVENTION:  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/10/028,072  
; CURRENT FILING DATE: 2001-12-19  
; PRIOR APPLICATION NUMBER: 60/049911  
; PRIOR FILING DATE: 1997-06-18  
; PRIOR APPLICATION NUMBER: 60/056974  
; PRIOR FILING DATE: 1997-08-26  
; PRIOR APPLICATION NUMBER: 60/059113  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059115  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059117  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059122  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059184  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059263  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/059352  
; PRIOR FILING DATE: 1997-09-19  
; PRIOR APPLICATION NUMBER: 60/059588  
; PRIOR FILING DATE: 1997-09-19  
; PRIOR APPLICATION NUMBER: 60/059836  
; PRIOR FILING DATE: 1997-09-24  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/062285  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/063287  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/062814  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/062816  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063045  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063082  
; PRIOR FILING DATE: 1997-10-31

```

; PRIOR APPLICATION NUMBER: 60/063127
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063327
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063329
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063550
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063561
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063704
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063733
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063735
; PRIOR FILING DATE: 1997-10-30
; PRIOR APPLICATION NUMBER: 60/063735
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063738
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064248
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/064480
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065846
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/066453
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066670
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/069212
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069278
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069334
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069694
; PRIOR FILING DATE: 1997-12-16
; PRIOR APPLICATION NUMBER: 60/072320
; PRIOR FILING DATE: 1998-01-23
; PRIOR APPLICATION NUMBER: 60/073612
; PRIOR FILING DATE: 1998-02-04
; PRIOR APPLICATION NUMBER: 60/074086
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/074092
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/074494
; PRIOR FILING DATE: 1998-02-27
; PRIOR APPLICATION NUMBER: 60/074963
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-02-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081695
; PRIOR FILING DATE: 1998-04-14
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081818

; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082999
; PRIOR FILING DATE: 1998-04-24
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085149
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/085697
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/086414
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/086430
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/0867106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-05-10
; PRIOR APPLICATION NUMBER: 60/088430
; PRIOR FILING DATE: 1998-05-11
; PRIOR APPLICATION NUMBER: 60/088741
; PRIOR FILING DATE: 1998-05-11
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-05-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-05-11
; PRIOR APPLICATION NUMBER: 60/0889532
; PRIOR FILING DATE: 1998-05-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-05-17
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-05-19
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-05-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-05-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-05-24
; PRIOR APPLICATION NUMBER: 60/090538
; PRIOR FILING DATE: 1998-05-24
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-05-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07

```

## Alignment Scores:

Pred. No.: 4.89e-30

Length: 290

Score: 563.50

Matches: 110

Conservative: 44

Percent Similarity: 61.85%

Best Local Similarity: 44.18%

Mismatches: 86

query Match: 38 65% Indels: 9 Gaps: 4

DB: US-09-598-982-20 (1-771) x US-10-028-072-222 (1-290)

QY 16 AGAACCTCGGGGTCAAGGAGCCCCAGGAGAACGAGTGGCCCTGGCAGGTGACCTGAGA 75  
 DB 34 ArgMetValGlyLgLyGinAspThrGlnGluGlyLgLyGlnProPrGlnValSerIleIn 53  
 QY 76 GTCCACGCCCATACTGGATGACTTCGCGGGCCTCCATCCACCCCCAGTGGGTG 135  
 DB 54 ArgAsnGly-----SerHisPheCysGlyLysLeuIleAlaGluGinTrpVal 70

QY 136 CTGACGCCG 195  
 DB 71 LeutIralAlaHisCysPhe--ArgAsnThrSerGluThrSerLeutYrgInValleu 89  
 QY 196 CTGGGGAGCACGACCCTACTAC-----CAGGACAGCTGCTCCGGTCAGG 246  
 DB 90 IeuglyIalaArgGlnLeuValGlnProGlyProHisAlaMetTyRalaArgValArgIn 109  
 QY 247 ATCATCGTGACCCACAGTCTACACGCCAGATCGGAGCGCAGATCGGCCCTGCTGGG 306  
 DB ::::: 110 ValGluSerAsnProLeuTyrgInGlyThrAlaSerSerAlaAspValAlaLeuValGlu 129

QY 307 CTGGAGCGCGCGCTGCGGGAGCGAGTCAGGATCTGGCCGGCATGTTGACAATGAT 426  
 DB 130 IeuglyIalaProValProPheThrAsnTyryleLeuProValCysLeuProAspPro 149

QY 367 GAGACCTCCCCCGGGATGCCGTGGTCACTGCTGGGCACTGCTGGGGCATGTTG 426  
 DB 150 ValIlePheGluThrGlyMetAsnCysTrpValThrGlyLgLyTrpGlySerProSerGluGlu 169

QY 427 GAGGCCCTCCACCCGCATTCTCTGAAAGCAGGTGAAGTCCCATAATGGAAACAC 486  
 DB 170 AspLeuLeuProGluProArgIleLengInlySleuAlaValProIleLeuAspThrPro 189

QY 487 ATTGTCGACGCAAATACACCTGGGCTACACGGAGACGAGTCGGCCATGCTGG 546  
 DB 190 LysCysAsnLeuLeutYryleLeuAspThrGluLpHeGlyTyrcInProLySthrIle 209

QY 547 GACGACTGCTGTCGGGG-----AACACGGAGGACTCATGCCAGGGCAGCTCC 600  
 DB 210 AsnAspMetLeuCysAlaIaglypheGluIgluGlyLysAspAlaCysLysGlyAspSer 229

QY 601 GGAGGGCCCTGGTGTGCAASSTGAAATGGCACCTGGCAGGGGGCTGGTCAGCTGG 660  
 DB 230 GlyGlyProLeuValGlyLgLyGinSerTrpLeuGlnAlaIgLyValleSerIpp 249

QY 661 GCGGAGGGCTGGCCACGCCAACGGCCCTGGCATCTACACCGGTGTCACACTTG 720  
 DB 250 GlyGlyLycySalaArgGlnAsnArgProGlyLysValTyrIleArgValThrAlaHis 269

QY 721 GACTGGATCACCACATATGTCGCCAA 747  
 DB 270 AsnTrpIleHisArgIleIleProLys 278

RESULT 5  
 US-10-040-655-7  
 ; Sequence 7, Application US/10040655  
 ; Patent No. US2020146805A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Andrade-Gordon, Patricia  
 ; APPLICANT: Darrow, Andrew  
 ; APPLICANT: Q, Jian-shen  
 ; TITLE OF INVENTION: DNA encoding the novel human serine  
 ; FILE REFERENCE: ORT-1032  
 ; CURRENT APPLICATION NUMBER: US/10/040, 655  
 ; CURRENT FILING DATE: 2002-01-07  
 ; NUMBER OF SEQ ID NOS: 11  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO: 7

DB: US-09-598-982-20 (1-771) x US-10-040-655-7 (1-290)

QY 16 AGAACCTCGGGGTCAAGGAGCCCCAGGAGAACGAGTGGCCCTGGCAGGTGACCTGAGA 75  
 DB 34 ArgMetValGlyLgLyGinAspThrGlnGluGlyLgLyGlnProPrGlnValSerIleIn 53  
 QY 76 GTCCACGCCCATACTGGATGACTTCGCGGGCCTCCATCCACCCCCAGTGGGTG 135  
 DB 54 ArgAsnGly-----SerHisPheCysGlyLysLeuIleAlaGluGinTrpVal 70

QY 136 CTGACGCCG 195  
 DB 71 LeutIralAlaHisCysPhe--ArgAsnThrSerGluThrSerLeutYrgInValleu 89  
 QY 196 CTGGGGAGCACGACTCTACTAC-----CAGGACAGCTGCTGGGTCAGG 246  
 DB 90 IeuglyIalaArgGlnLeuValGlnProGlyProHisAlaMetTyRalaArgValArgIn 109  
 QY 247 ATCATCGTGACCCACAGTCTACACGCCAGATCGGAGCGCAGTGGCCCTGCTGGG 306  
 DB 110 ValGluSerAsnProLeuTyrgInGlyLysLeuIleAlaIgLyValleIle 129

QY 307 CTGGAGCGCGCGCTGCGGGAGCGAGTCAGGATCTGGCCGGCATGTTGACAATGAT 426  
 DB 130 IeuglyIalaProValProPheThrAsnTyryleLeuProValCysLeuProAspPro 149

QY 367 GAGACTCTCCCCCGGGATGCCGTGGTCACTGCTGGGCACTGCTGGGGCATGTTG 426  
 DB 150 ValIlePheGluThrGlyMetAsnCysTrpValThrGlyLgLySerProSerGluGlu 169

QY 427 GAGCCCTCCACCCGCATTCTCTGAAAGCAGGTGAAGTCCCATAATGGAAACAC 486  
 DB 170 AspLeuLeuProGluProArgIleLengInlySleuAlaValProIleLeuAspThrPro 189

QY 487 ATTGTCGACGCAAATACACCTGGGCTACACGGAGACGAGTCGGCCATGCCCGT 546  
 DB 190 LysCysAsnLeuLeutYryleLeuAspThrGluLpHeGlyTyrcInProLySthrIle 209

QY 547 GACGATGCTGTCGGCCSGG-----AACACCGGAGGACTCATGCCAGGGCAGCTCC 600  
 DB 210 AsnAspMetLeuCysAlaIaglypheGluIgluGlyLysAspAlaCysLysGlyAspSer 229

QY 601 GGAGGGCCCTGGTGTGCAAGGTGATGGCACCTGGCTGAGGGGGCTGGTCAGCTGG 660  
 DB 230 GlyGlyProLeuValGlyLgLyGinSerTrpLeuGlnAlaIgLyValleSerIpp 249

QY 661 GCGGAGGGCTGGCCACGCCAACGGCCCTGGCATCTACACCGGTGTCACACTTG 720  
 DB 250 GlyGlyLycySalaArgGlnAsnArgProGlyLysValTyrIleArgValThrAlaHis 269

QY 721 GACTGGATCACCACATATGTCGCCAA 747  
 DB 270 AsnTrpIleHisArgIleIleProLys 278

RESULT 6  
 US-10-041-006A-9  
 ; Sequence 9, Application US/10041006A  
 ; Patent No. US2020168734A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Andrade-Gordon, Patricia

APPLICANT: Darrow, Andrew  
 APPLICANT: Oi, Jian-shen  
 TITLE OF INVENTION: DNA encoding the novel human serine protease T  
 FILE REFERENCE: ORI-1032  
 CURRENT APPLICATION NUMBER: US/10/041,006A  
 CURRENT FILING DATE: 2002-01-07  
 NUMBER OF SEQ ID NOS: 11  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO: 9  
 LENGTH: 315  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Description of Artificial Sequence: Fusion Protein  
 OTHER INFORMATION: of Protease T in a zymogen activation construct  
 US-10-041-006A-9

Alignment Scores:  
 Pred. No.: 1.9e-29  
 Score: 554.50  
 Percent Similarity: 60.96%  
 Best Local Similarity: 43.43%  
 Query Match: 38.03%  
 DB: 9

Length: 315  
 Matches: 109  
 Conservative: 44  
 Mismatches: 89  
 Indels: 9  
 Gaps: 4

US-09-598-982-20 (1-771) x US-10-041-006A-9 (1-315)

QY 10 GAGAAAGATCGTCGGGCTCAGGAGCCCCAGGAGCAACTGGCCCTGGGCTGGCAGGTGAGC 69  
 Db 49 AspAspLysIleValGlyGlyTyrAlaLeuGluGluGlyGluTrpProTrpGlnValSer 68

QY 70 CTGAGAGTCACGGCCATACTGGATCACTrCTGGGGGSCCTCATCACCCCCAG 129  
 Db 69 IleGlnArgAsnGly-----SerHisPheCysGlyGlySerIleUileAlaGluGin 85

QY 130 TGGGTGTGACGGGGGGGGGGTGGGGGACGGACGCTCAAGATCTGGCCGCCCTGAGG 189  
 Db 86 TrpValLeuThrAlaAlaLahisCysPhe--ArgAsnThrSerGluThrSerLeuTyrgln 104

QY 190 GTGCAACTCGGGGAGCAGCTACTAC-----CAGGACAGCTGCTGCCGTC 240  
 Db 105 ValLeuLeuGlyAlaArgGlnLeuValGlnProGlyProHisAlaMetTyraVal 124

QY 241 AGCAGGATCATCTGGACCCACAGTCTACACCGCCAGATGGACGGGACATGCCCTG 300  
 Db 125 ArgGlnValGluUserAsnProLeuTyrgInGlyThrAlaSerSerAlaAspValAlaLeu 144

QY 301 CTGGAGCTGGAGGCCGGTGAGGCTCTCACGCCACCTCACACGGTCACCTGCCCT 360  
 Db 145 ValGluLeuGluAlaProValProPheThrAsnProLeuProAlaCysLeuProAsp 164

QY 361 GCCTCAAGACCTCCCCCGGGATGCCGCTGGGCTACTGGCTGGGCACTGGAC 420  
 Db 165 ProSerValIlePheGluThrGlyMetAsnCysTrpValIleGlyTrpGlySerProSer 184

QY 421 AATGATGAGGCCCTCCACGCCATTCTGAGACAGGAGAACGCTCCATAATGGA 480  
 Db 185 GluGluLeuAspLeuLeuProGluProArgIleLeuGlyIleLeuAlaValProIleLeuAsp 204

QY 481 ACCACATTTGACGGAAATAACACCTTGCGGCCACACGGAGAGCGACGCCGCGATC 540  
 Db 205 ThrProLysCysAsnLeuLeuTyrSerLysAspThrGluPheGlyTyrGlnProLysThr 224

QY 541 GTCCGAGACGACATGCTGTCGTCGGCCTACACGGAGAGCGACGCCGAGGC 594  
 Db 225 IleLysAsnAspMetIleLeuCysLagIgYPhyGluGluGlyLysLysAspAlaLysIleGly 244

QY 595 GACTCGGAGGCCCTGGTGCAGGTGAATGGCACCTGCTGCAGGGGGCTGGTGC 654  
 Db 245 AspSerGlyIgYProLeuValCysLeuValGlyIleSerTrpLeuGlnAlaIleValle 264

QY 655 AGCTGGGGGAGGGCTCTGCCAGGCCAACCGCCCTGGCTACACCCGGTGCACCTAC 714

RESULT 7  
 US-10-040-655-9  
 Sequence 9, Application US/10040655  
 Patent No. US20020146805A1  
 GENERAL INFORMATION  
 APPLICANT: Darrow, Andrew  
 APPLICANT: Andrade-Gordon, Patricia  
 APPLICANT: Darrow, Andrew  
 APPLICANT: Oi, Jian-shen  
 TITLE OF INVENTION: DNA encoding the novel human serine protease T  
 FILE REFERENCE: ORI-1032  
 CURRENT APPLICATION NUMBER: US/10/040,655  
 CURRENT FILING DATE: 2002-01-07  
 NUMBER OF SEQ ID NOS: 11  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO: 9  
 LENGTH: 315  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Description of Artificial Sequence: Fusion Protein  
 OTHER INFORMATION: of Protease T in a zymogen activation construct  
 US-10-040-655-9

Alignment Scores:  
 Pred. No.: 1.9e-29  
 Score: 554.50  
 Percent Similarity: 60.96%  
 Best Local Similarity: 43.43%  
 Query Match: 38.03%  
 DB: 12

Length: 315  
 Matches: 109  
 Conservative: 44  
 Mismatches: 89  
 Indels: 9  
 Gaps: 4

US-09-598-982-20 (1-771) x US-10-040-655-9 (1-315)

QY 10 GAGAAAGATCGTCGGGCTCAGGAGCCCCAGGAGCAACTGGCCCTGGGCTGGCAGGTGAGC 69  
 Db 49 AspAspLysIleValGlyGlyTyrAlaLeuGluGluGlyGluTrpProTrpGlnValSer 68

QY 70 CTGAGAGTCACGGCCATACTGGATCACTrCTGGGGGSCCTCATCACCCCCAG 129  
 Db 69 IleGlnArgAsnGly-----SerHisPheCysGlyGlySerIleUileAlaGluGin 85

QY 130 TGGGTGTGACGGGGGGGGGGTGGGGGACGGACGCTCAAGATCTGGCCGCCCTGAGG 189  
 Db 86 TrpValLeuThrAlaAlaLahisCysPhe--ArgAsnThrSerGluThrSerLeuTyrgln 104

QY 190 GTGCAACTCGGGGAGCAGCTACTAC-----CAGGACAGCTGCTGCCGTC 240  
 Db 105 ValLeuLeuGlyAlaArgGlnLeuValGlnProGlyProHisAlaMetTyraVal 124

QY 241 AGCAGGATCATCTGGACCCACAGTCTACACCGCCAGATGGACGGGACATGCCCTG 300  
 Db 125 ArgGlnValGluUserAsnProLeuTyrgInGlyThrAlaSerSerAlaAspValAlaLeu 144

QY 301 CTGGAGCTGGAGGCCGGTGAGGCTCTCACGCCACCTCACACGGTCACCTGCCCT 360  
 Db 145 ValGluLeuGluAlaProValProPheThrAsnProLeuProAlaCysLeuProAsp 164

QY 361 GCCTCAAGACCTCCCCCGGGATGCCGCTGGGCTACTGGCTGGGCACTGGAC 420  
 Db 165 ProSerValIlePheGluThrGlyMetAsnCysTrpValIleGlyTrpGlySerProSer 184

QY 421 AATGATGAGGCCCTCCACGCCATTCTGAGACAGGAGAACGCTCCATAATGGA 480  
 Db 185 GluGluLeuAspLeuLeuProGluProArgIleLeuGlyIleLeuAlaValProIleLeuAsp 204

QY 481 ACCACATTTGACGGAAATAACACCTTGCGGCCACACGGAGAGCGACGCCGAGGC 540  
 Db 205 ThrProLysCysAsnLeuLeuTyrSerLysAspThrGluPheGlyTyrGlnProLysThr 224

QY 541 GTCCGAGACGACATGCTGTCGTCGGCCTACACGGAGAGCGACGCCGAGGC 594  
 Db 225 IleLysAsnAspMetIleLeuCysLagIgYPhyGluGluGlyLysLysAspAlaLysIleGly 244

QY 595 GACTCGGAGGCCCTGGTGCAGGTGAATGGCACCTGCTGCAGGGGGCTGGTGC 654  
 Db 245 AspSerGlyIgYProLeuValCysLeuValGlyIleSerTrpLeuGlnAlaIleValle 264

QY 655 AGCTGGGGGAGGGCTCTGCCAGGCCAACCGCCCTGGCTACACCCGGTGCACCTAC 714



DB: 12 Gaps: 4

QY 127 CAGTGGGTGTGACCGCCGGCTGGGGACTGTCANGATCTGGCCCTC 186  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 70 GlnTrpValLeuThrAlaLahisCysPhe---ProArgArgAlaLeuProAlaIglnYr 88  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 187 AGGGTGAACACTGCAGGAGCACCTC-----TACTACCGACGCTGCG 237  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 89 ArgValArgLargIleuGlyValAlaArgLeuIglnYrSerThrSerProArgThrLeuSerValPro 108  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 238 GTCAcAGAGATTCACTCGTGACCCACAGTCTACACGCCAGAHCAGGAGGGACATCCC 297  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 109 ValArgArgValLeuLeuProProAspTyrsSerGluAspGlyAlaArgGlyAspLeuAla 128  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 358 CCTGCCTAGAGACCTCCCCGGATGCCGTACTGGCTGGGTACTGGCTGGGGATG 417  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 298 CTGGCTGGACTCTGGAGGGGGTGAAGGCTCCAGGCCACTCCACCGCTACCTGCC 357  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 418 GACAATGAGCAGGCCCTCCACCGCATTCTCTGAACAGGGAAGSTCCCATATG 477  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 129 LeuLeuGlyLeuIglnLeuArgArgProvalProleuSerAlaArgValGlyAspLeuPro 148  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 538 ATCGTCGGTAGACGACATGCTGTGCCGG-----AACACCGGAGGACTATGCCAG 591  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 209 IleValLeuProGlySerLeuCysAlaGlyTyroProGlyIleGlyVal 228  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 592 GGCGACTCCGGAGGGGCCCTGGTGAAGGTGAATGGACCTGGCTGAGGGCGCG 651  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 189 AspSerArgThrCysAspGlyLeuTyrrHisValGlyAlaAspGlyAlaIglnYr 208  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 538 ATCGTCGGTAGACGACATGCTGTGCCGG-----AACACCGGAGGACTATGCCAG 591  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 249 ValSerTrpGlyIleGlyCysAlaLeuProAsnArgProGlyValTyrrHisValAla 268  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 712 TACTACTGGACTGGATC 729  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 269 ThrTyrSerProTrpIle 274  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

---

RESULT 10

US-10-042-091A-7

; Sequence 7, Application US/10042091A

; Patent No. US20020142447A1

; GENERAL INFORMATION:

; APPLICANT: Darrow, Andrew

; APPLICANT: Andrade-Gordon, Patricia

; APPLICANT: Qi, Jenson

; TITLE OF INVENTION: DNA Encoding the Human Serine

; TITLE OF INVENTION: Protease EOS

; FILE REFERENCE: ORT-1031

; CURRENT APPLICATION NUMBER: US/10/042, 091A

; CURRENT FILING DATE: 2002-01-08

; PRIOR APPLICATION NUMBER: US/09/387, 375

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: PatentIn Ver. 2.0

; LENGTH: 284

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-10-042-091A-7

---

Alignment Scores:

Pred. No.: 2.86e-28 Length: 284

Score: 536.50 Matches: 112

Percent Similarity: 58.54% Conservative: 312

Best Local Similarity: 45.53% Mismatches: 93

Query Match: 36.80% Indels: 9

---

RESULT 11

US-10-041-400A-9

; Sequence 9, Application US/10041400A

; Patent No. US20011095A1

; GENERAL INFORMATION:

; APPLICANT: Darrow, Andrew

; APPLICANT: Andrade-Gordon, Patricia

; APPLICANT: Qi, Jenson

; TITLE OF INVENTION: DNA Encoding the Human Serine

; TITLE OF INVENTION: Protease EOS

; FILE REFERENCE: ORT-1031

; CURRENT APPLICATION NUMBER: US/10/041, 400A

; CURRENT FILING DATE: 2002-01-08

; PRIOR APPLICATION NUMBER: US/09/387, 375

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: PatentIn Ver. 2.0





QY 241 AGCAGGATCATCGTCACCCACACTCTACACCCCCAGATCGGAGCGACATGCCCTG 300  
 ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 118 LysAspIleLeuSerProHisProSerTyroLeuIleGluGlySerGlnGlyAspIleAlaLeu 137  
 QY 301 CTGGAGCTGAGGGGCCGTGAGGCTGAGGCACTGAGGTCACAGGTGACCCGCCCT 360  
 ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 138 LeuGlnLeuSerArgProIleThrPheSerArgTyrIleArgProIleCysLeuProAla 157  
 QY 361 GGCCTCAGAGACTTCCCCGGGATGCGCTGCGGTACTGGTGGGGATGGGCA 420  
 ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 158 AlaAsnAlaSerPheProSnglyIleHisCysThrValThrGlyTrpGlyHsvalAla 177  
 QY 481 AACCCATATGAGGCAAATAACCCACCTGGCCCTAACGGGAGACCAAGTGGCAC 540  
 ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 198 ArgGluThrCysAsnCysIleUtyrasNleAspAla---LysProGluGluProHisPhe 216  
 QY 541 GTCCGGTGACACAGCTGCTGCGGGAAACACCCGG-----AGGACTCATGCCAGGGC 594  
 ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 217 ValGlnGlyAspMetValCysAlaGlyIleUtyrValGlyUglyIlyAspAlaCys 236  
 QY 595 GACTCGGAGGGCCCTGGTGGAAAGGTAATGCCACCTGGCTGAGGGCTGGTC 654  
 ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 237 AspSerGlyIleProLeuSerCysProValGluGlyIleUtyrIleThrGlyIleVal 256  
 QY 655 AGCTGGGGCAGGGCTGTGCCAACGCCAACGGGCTGGCATCTAACCCGGTGTAC 714  
 ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 257 SerItpGlyAspAlaCysGlyIlaArgAshargProGlyValYrThrIleAlaSerSer 276  
 QY 715 TACTGGACGGATCACCCTAATGTC 741  
 ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 277 TyrAlaSerItpGlyIleSerItpGlyIleVal 285

RESULT 15

US-10-067-761-19

; Sequence 19, Application US/10067761

; Publication No. US20020197701A1

; GENERAL INFORMATION:

; APPLICANT: Shi et al.

; TITLE OF INVENTION: Serine protease Polynucleotides, Polypeptides, and Antibodies

; FILE REFERENCE: PT00524

; CURRENT APPLICATION NUMBER: US10/067,761

; CURRENT FILING DATE: 2002-02-08

; PRIOR APPLICATION NUMBER: 09/804,156

; PRIOR FILING DATE: 2001-01-13

; PRIOR APPLICATION NUMBER: 60/189,025

; PRIOR FILING DATE: 2000-03-14

; NUMBER OF SEQ ID NOS: 38

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO: 19

; LENGTH: 238

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-10-067-761-19

Alignment Scores:

Pred. No.:	3.53e-26	Length:	238
Score:	504.50	Matches:	106
Percent Similarity:	59.73%	Conservative:	26
Best Local Similarity:	47.96%	Mismatches:	76
Query Match:	34.60%	Indels:	13
DB:	9	Gaps:	6

US-09-598-982-20 (1-771) x US-10-067-761-19 (1-238)

QY 16 AGAATCTGTCGGGGCAGGAGCAGTCAGGCTGGCTGGAGGCTGGAGCTGAGA 75  
 ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 29 ArgIleValGlyGlyHisAlaAlaProAlaGlyIalaTrpProTrpGlnAlaSerLeuArg 48

QY 76 GTCCACGGCCCTACTGGATGCACTCTGGGGCTCCCTCATCCACCCCCAGGGTG 135  
 :::: :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 49 LeuArgArg-----valHisValCysGlyIleSerIleLeuSerProGlnTrpAl 65  
 QY 136 CTGACCGCGGGCTGCGGGACGGACGTCAAGGACTGGTGGCCCTCAGGGTCAA 195  
 ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 66 LeutInAlaAlaAlaHisCysPheSerGlySerLeuAsnSer--SerAspTyrgInValHis 84  
 QY 196 CTGGGGACAGCACCTCTACTACAGGACCCACGTCGCTCCGGCAGAGTCATCCG 255  
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 85 LeuGlyIleUtyrGluIleThrLeuSerProHisPheSerThrValArgGlnIleIleLeu 104  
 QY 256 CACCCACAGTCTACACCGCCAGATCGGA----GCGGACATGCCCTGCTGGAGCTG 309  
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 105 HisSerSer---ProSerGlyIleProGlyIleSerGlyAspIleAlaLeuValGluLeu 123  
 QY 310 GAGGAGCGGTGAAGGTCMCCAGCACGGTACCGGCTACGCCCTGCCCCTGCTCAGAG 369  
 ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 124 SerValProValIleLeuSerArgIleLeuProValCysIleUproGluAlaSerAsp 143  
 QY 370 ACCTTCCCCCGGGATGCGCTGCGGCTACAGCTGGGGATGTGGACATGATGATG 429  
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 144 AspHeCysProGlyIleArgCysTrpValThrGlyIleTrpGlyIleVal 163  
 QY 430 CGCCCTCCACGCCATTCTCTGAGCACGGTGAGGTCCTCATATGAAACACATT 489  
 ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 164 ProLeuProProProTyrsSerIleArgGluUtyrValAspThrGluThr 183  
 QY 490 TGTGACGCAAATAACCAACCTGGCCCTACAGGAGACGAGCTGGCATCGTGC 549  
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 184 CysArgArgAsp-----TyrProGlyProGlyIleSerIleLeuGlnPro 198  
 QY 550 GACATGCTGTTGCGGAAACACCGGAGGACTCATGCCAGGGGACTCCGGGGGCC 609  
 ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 199 AspMetLeuCysAlaArgGlyPro---GlyAspAlaCysGlnAspAspSerGlyGlyPro 217  
 QY 610 CTGGTGTGCCAGGTCATGCCACCTGGCTCCAGGGGGCTGGCAGCGGGGAGGCC 669  
 ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 218 IleUtyrCysIleUtyrValAsnGlyIalaTrpValGlnAlaGlyIleUtyr 237  
 QY 670 TGT 672  
 Db 238 Cys 238

Search completed: January 31, 2003, 07:16:39  
 Job time : 45 secs



N/0

OM nucleic - protein search, using frame\_plus\_n2p model

Run on : January 31, 2003, 06:51:34 ; Search time 25.5 Seconds  
 (without alignments) 5813.305 Million cell updates/sec

Title: US-09-598-982-20

Perfect score: 1458

Sequence: 1. 999ccctcgagaaaaaat.....cgtgaagcgccgcgtcg 771

Scoring table: BLOSUM62

Xgapop	Ygapext	Xgapext	Ygapop	Fgapext	Delop
10.0	0.5	0.5	10.0	7.0	6.0 , Delext 7.0

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 566448

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```
-MODEL=frame+ n2p model -DEY=x1h
-Q=/gn2/1/USP10_spool/US9598982/runat_27012003_073717_19581/app_query.fasta_1.967
-DB=PIR_73 -QMM=fastan -SUFIX=n2p.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIXX=blosum62 -TRANS=human40.cgi -LIST=45
-DOCIGN=200 -THR_SCORPCT=THR_MAX100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT_PTO -NORMEXT -HEAPSIZE=20000000000
-USER=US9598982 @CGN_1_1_15 @runat_27012003_073717_19581 -NCPU=3
-NO_XLPPX -NO_MAP -LARGEQUERY -NEGSCORES=0 -WAIT=1 -LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPEXT=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database :

```
PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

RESULT 1

B35663

trypase (EC 3.4.21.59) II precursor - human

N;Alternate names: trypase beta

C;Species: Homo sapiens (man)

C;Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 21-Jul-2000

C;Accession: B35663; R37193; I59473

R;Vanderslice, P.; Ballinger, S.M.; Tam, E.K.; Goldstein, S.M.; Craik, C.S.; Caughey, Proc. Natl. Acad. Sci. U.S.A. 87, 3811-3815, 1990

A;Title: Human mast cell trypsin: multiple cDNAs and genes reveal a multigene serine

A;Reference number: A35663; MUID:90251647; PMID:2187193

A;Molecule type: mRNA; DNA

A;Residues: 1-275 <VAN>

A;Cross-references: GB:33492; NID:9339982; PIDN:AAA36779\_1; PID:9339983

A;Note: residues 2-275 are derived from mRNA; residue one was inferred from the genome

R;Miller, J.S.; Moxley, G.; Schwartz, L.B.

J;Clin. Invest. 86, 865-870, 1990

A;Title: Cloning and characterization of a second complementary DNA for human trypas A;Reference number: A37193; MUID:90369005; PMID:2203827

A;Accession: A37193

A;Molecule type: mRNA

A;Residues: 1-275 <MLT>

A;Cross-references: GB:M37488; NID:9179583; PIDN:AAA51843\_1; PID:9179584

R;Blom, T.; Hallman, L.

Scard, J. Immunol. 37, 203-208, 1993

A;Title: Characterization of a trypase mRNA expressed in the human basophil cell line

A;Reference number: I59473; MUID:93166209; PMID:8434231

A;Accession: I59473

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-275 <RES>

A;Cross-references: GB:55551; NID:g26566; PIDN:AA13876\_1; PID:94261576

A;Experiments: basophil cell line KU812

C;Genetics:

A;Gene: GDB:TPS1

A;Cross-references: GDB:125890; OMIM:191080

A;Map position: 16pter-16qter

Copyright (c) 1993 - 2003 Compugen Ltd.

GenCore version 5.1.3

Score	Match	Length	DB	ID
15	451	30.9	638	1 KOHUP
16	448	30.7	455	2 A61545
17	447	30.7	638	1 KQMSPL
18	46.5	30.6	812	1 PLBO
19	44.5	30.5	638	1 KORTPL
20	44.5	30.5	271	1 KURT2
21	44.3	30.4	245	1 KYBOB
22	44.1	30.2	460	2 B61545
23	43.9	30.1	269	2 B26823
24	43.7	30.0	1524	2 T30337
25	43.5	29.5	271	2 A25528
26	42.9	29.4	810	2 PIHU
27	42.8	29.4	810	2 B30848
28	42.7	29.3	263	1 A21195
29	42.7	29.3	626	1 ELPG
30	42.6	29.2	263	2 A31299
31	42.2	28.9	263	1 KYRTB
32	42.0	28.8	812	1 PLMS
33	42.0	28.8	786	1 A47547
34	41.9	28.7	245	1 KYBOA
35	41.9	28.7	558	1 JC5878
36	41.9	28.7	790	1 PLPG
37	41.6	28.6	246	2 B25528
38	41.4	28.4	246	1 TRRT2
39	41.4	28.4	264	2 I38136
40	41.3	28.4	269	2 C26823
41	41.3	28.4	367	1 JE0104
42	41.2	28.3	246	1 TRT1
43	41.2	28.3	258	2 S70439
44	41.2	28.3	267	4 A56615
45	41.1	28.2	1420	2 A32869

plasma kallikrein

plasmin (EC 3.4.21

plasma kallikrein

plasmin (EC 3.4.21

pancreatic elastase

chymotrypsin (EC 3

chymotrypsin (EC 3

pancreatic elastase

pancreatic elastase

pancreatic elastase

pancreatic elastase

plasmin (EC 3.4.21

serine protease

chymotrypsin (EC 3

plasma hyaluronan

plasmin (EC 3.4.21

chymotrypsin (EC 3.4.21

chymotrypsin-like

pancreatic elastase

testicular serine

trypsin (EC 3.4.21

pancreatic elastase

probable pancreatic

apolipoprotein(a)

C; Superfamily: trypsin; trypsin homology  
 C; Keywords: hydrolase; serine proteinase; zymogen  
 C; I-21-Domain: signal sequence #status predicted <SIG>  
 F; 22-30/Domain: activation peptide #status predicted <ACT>  
 F; 31-275/Product: tryptase I #status predicted <MAT>  
 F; 74, 121, 224/Active site: trypsin homology <TRY>  
 A; title: Human mast cell trypsin: multiple cDNAs and genes reveal a multigene serineel  
 A; Reference number: A35863; MUID:90251647; PMID:218793  
 A; Accession: A35863  
 A; Molecule type: DNA  
 A; Residues: 1-275 <VAN>  
 A; Cross-references: GB:M33491  
 A; Cross-references: GB:N33494; NID:93927804; PIDN: AAC83172.1; PID:9339977  
 A; Accession: D35863  
 A; Molecule type: mRNA  
 A; Residues: 1-275 <VA2>  
 A; Cross-references: GB:M33491  
 R; Butterfield, J.H.; Weiler, D.A.; Hunt, L.W.; Wynn, S.R.; Roche, P.C.  
 J. Leukoc. Biol., 47, 409-419, 1990  
 A; Purification of tryptase from a human mast cell line.  
 A; Reference number: A60939; MUID:90241210; PMID:2110591  
 A; Accession: A60939  
 A; Molecule type: protein  
 A; Residues: 31-38, P, 40-41, X, 43, T, 45-48, X, 50 <BUT>  
 A; Experimental source: mast cell  
 A; Note: 44-GLY was also found  
 R; Cromlish, J.A.; Seidal, N.G.; Marcinkiewicz, M.; Hamelin, J.; Johnson, D.A.; Chrestie  
 J. Biol. Chem., 262, 1363-1373, 1987  
 A; Title: Human pituitary trypsin: molecular forms, NH-2-terminal sequence, immunocyt  
 A; Reference number: A39326; MUID:87109258; PMID:3543004  
 A; Molecule type: protein  
 A; Residues: 31-38 <CRO>  
 A; Experimental source: pituitary  
 C; Genetics:  
 A; Introns: 21/1; 78/2; 177/1; 221/3  
 C; Superfamily: trypsin; trypsin homology  
 C; Keywords: hydrolase; serine proteinase; zymogen  
 F; 22-30/Domain: activation peptide #status predicted <SIG>  
 F; 31-275/Product: tryptase I #status predicted <ACT>  
 F; 74, 121, 224/Active site: His, Asp, Ser #status predicted  
 A; title: Human mast cell trypsin: multiple cDNAs and genes reveal a multigene serineel  
 A; Reference number: A35863; MUID:90251647; PMID:218793  
 A; Accession: A35863  
 A; Molecule type: DNA  
 A; Residues: 1-275 <VAN>  
 A; Cross-references: GB:M33491  
 A; Cross-references: GB:N33494; NID:93927804; PIDN: AAC83172.1; PID:9339977  
 A; Accession: D35863  
 A; Molecule type: mRNA  
 A; Residues: 1-275 <VA2>  
 A; Cross-references: GB:M33491  
 R; Butterfield, J.H.; Weiler, D.A.; Hunt, L.W.; Wynn, S.R.; Roche, P.C.  
 J. Leukoc. Biol., 47, 409-419, 1990  
 A; Purification of tryptase from a human mast cell line.  
 A; Reference number: A60939; MUID:90241210; PMID:2110591  
 A; Accession: A60939  
 A; Molecule type: protein  
 A; Residues: 31-38, P, 40-41, X, 43, T, 45-48, X, 50 <BUT>  
 A; Experimental source: mast cell  
 A; Note: 44-GLY was also found  
 R; Cromlish, J.A.; Seidal, N.G.; Marcinkiewicz, M.; Hamelin, J.; Johnson, D.A.; Chrestie  
 J. Biol. Chem., 262, 1363-1373, 1987  
 A; Title: Human pituitary trypsin: molecular forms, NH-2-terminal sequence, immunocyt  
 A; Reference number: A39326; MUID:87109258; PMID:3543004  
 A; Molecule type: protein  
 A; Residues: 31-38 <CRO>  
 A; Experimental source: pituitary  
 C; Genetics:  
 A; Introns: 21/1; 78/2; 177/1; 221/3  
 C; Superfamily: trypsin; trypsin homology  
 C; Keywords: hydrolase; serine proteinase; zymogen  
 F; 22-30/Domain: activation peptide #status predicted <SIG>  
 F; 31-275/Product: tryptase I #status predicted <ACT>  
 F; 74, 121, 224/Active site: His, Asp, Ser #status predicted  
 A; title: Human mast cell trypsin: multiple cDNAs and genes reveal a multigene serineel  
 A; Reference number: A35863; MUID:90251647; PMID:218793  
 A; Accession: A35863  
 A; Molecule type: DNA  
 A; Residues: 1-275 <VAN>  
 A; Cross-references: GB:M33491  
 A; Cross-references: GB:N33494; NID:93927804; PIDN: AAC83172.1; PID:9339977  
 A; Accession: D35863  
 A; Molecule type: mRNA  
 A; Residues: 1-275 <VA2>  
 A; Cross-references: GB:M33491  
 R; Butterfield, J.H.; Weiler, D.A.; Hunt, L.W.; Wynn, S.R.; Roche, P.C.  
 J. Leukoc. Biol., 47, 409-419, 1990  
 A; Purification of tryptase from a human mast cell line.  
 A; Reference number: A60939; MUID:90241210; PMID:2110591  
 A; Accession: A60939  
 A; Molecule type: protein  
 A; Residues: 31-38, P, 40-41, X, 43, T, 45-48, X, 50 <BUT>  
 A; Experimental source: mast cell  
 A; Note: 44-GLY was also found  
 R; Cromlish, J.A.; Seidal, N.G.; Marcinkiewicz, M.; Hamelin, J.; Johnson, D.A.; Chrestie  
 J. Biol. Chem., 262, 1363-1373, 1987  
 A; Title: Human pituitary trypsin: molecular forms, NH-2-terminal sequence, immunocyt  
 A; Reference number: A39326; MUID:87109258; PMID:3543004  
 A; Molecule type: protein  
 A; Residues: 31-38 <CRO>  
 A; Experimental source: pituitary  
 C; Genetics:  
 A; Introns: 21/1; 78/2; 177/1; 221/3  
 C; Superfamily: trypsin; trypsin homology  
 C; Keywords: hydrolase; serine proteinase; zymogen  
 F; 22-30/Domain: activation peptide #status predicted <SIG>  
 F; 31-275/Product: tryptase I #status predicted <ACT>  
 F; 74, 121, 224/Active site: His, Asp, Ser #status predicted  
 A; title: Human mast cell trypsin: multiple cDNAs and genes reveal a multigene serineel  
 A; Reference number: A35863; MUID:90251647; PMID:218793  
 A; Accession: A35863  
 A; Molecule type: DNA  
 A; Residues: 1-275 <VAN>  
 A; Cross-references: GB:M33491  
 A; Cross-references: GB:N33494; NID:93927804; PIDN: AAC83172.1; PID:9339977  
 A; Accession: D35863  
 A; Molecule type: mRNA  
 A; Residues: 1-275 <VA2>  
 A; Cross-references: GB:M33491  
 R; Butterfield, J.H.; Weiler, D.A.; Hunt, L.W.; Wynn, S.R.; Roche, P.C.  
 J. Leukoc. Biol., 47, 409-419, 1990  
 A; Purification of tryptase from a human mast cell line.  
 A; Reference number: A60939; MUID:90241210; PMID:2110591  
 A; Accession: A60939  
 A; Molecule type: protein  
 A; Residues: 31-38, P, 40-41, X, 43, T, 45-48, X, 50 <BUT>  
 A; Experimental source: mast cell  
 A; Note: 44-GLY was also found  
 R; Cromlish, J.A.; Seidal, N.G.; Marcinkiewicz, M.; Hamelin, J.; Johnson, D.A.; Chrestie  
 J. Biol. Chem., 262, 1363-1373, 1987  
 A; Title: Human pituitary trypsin: molecular forms, NH-2-terminal sequence, immunocyt  
 A; Reference number: A39326; MUID:87109258; PMID:3543004  
 A; Molecule type: protein  
 A; Residues: 31-38 <CRO>  
 A; Experimental source: pituitary  
 C; Genetics:  
 A; Introns: 21/1; 78/2; 177/1; 221/3  
 C; Superfamily: trypsin; trypsin homology  
 C; Keywords: hydrolase; serine proteinase; zymogen  
 F; 22-30/Domain: activation peptide #status predicted <SIG>  
 F; 31-275/Product: tryptase I #status predicted <ACT>  
 F; 74, 121, 224/Active site: His, Asp, Ser #status predicted  
 A; title: Human mast cell trypsin: multiple cDNAs and genes reveal a multigene serineel  
 A; Reference number: A35863; MUID:90251647; PMID:218793  
 A; Accession: A35863  
 A; Molecule type: DNA  
 A; Residues: 1-275 <VAN>  
 A; Cross-references: GB:M33491  
 A; Cross-references: GB:N33494; NID:93927804; PIDN: AAC83172.1; PID:9339977  
 A; Accession: D35863  
 A; Molecule type: mRNA  
 A; Residues: 1-275 <VA2>  
 A; Cross-references: GB:M33491  
 R; Butterfield, J.H.; Weiler, D.A.; Hunt, L.W.; Wynn, S.R.; Roche, P.C.  
 J. Leukoc. Biol., 47, 409-419, 1990  
 A; Purification of tryptase from a human mast cell line.  
 A; Reference number: A60939; MUID:90241210; PMID:2110591  
 A; Accession: A60939  
 A; Molecule type: protein  
 A; Residues: 31-38, P, 40-41, X, 43, T, 45-48, X, 50 <BUT>  
 A; Experimental source: mast cell  
 A; Note: 44-GLY was also found  
 R; Cromlish, J.A.; Seidal, N.G.; Marcinkiewicz, M.; Hamelin, J.; Johnson, D.A.; Chrestie  
 J. Biol. Chem., 262, 1363-1373, 1987  
 A; Title: Human pituitary trypsin: molecular forms, NH-2-terminal sequence, immunocyt  
 A; Reference number: A39326; MUID:87109258; PMID:3543004  
 A; Molecule type: protein  
 A; Residues: 31-38 <CRO>

Db 131 ValAsnValSerSerHisValHisthValThrLeuProProLaserGluThrPhePro 150  
 QY ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 CCGGGGAGGCCGCGCTGGTCACTGGCTGGGGCGATGGGACATGATGAGGCCCTCCA 438  
 Db 151 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAsnAspGluArgLeuPro 170  
 QY ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 439 CCGCCATTCTCTGAAACAGGTGAAGGTCCTAATGGAAACACACATWGTGACCA 498  
 Db 171 ProProHeProLeuLysGlnValLysValProLeuMetGluAsnHisIleCysAspAla 190  
 QY ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 499 AAATACCAACTTGCGCTACAGGGAGCACGAGCTCCCATCTGCCCTGAGGTCAGT 558  
 Db 191 LYSThrHisLeuGlyAlaTyrrhGlyAspAspValAргLeuValArgLysAspMetIeu 210  
 QY ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 559 TGRGCCGCGAACCCGAGGGACTCTGCCAGGGACTCGGGGACCTGGGCCCTGGGCC 618  
 Db 211 CysAlaGlyAsnThrArgAspSerCysGlyLysGlyProAspValAspLeuValCys 230  
 QY ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 619 AGGGTGAATGGCACCTGGCTGGGGGGGGCTGGCTAGCTGGGGGGCTGCCCC 678  
 Db 231 LysValAsnGlyThrTrpLeuGlnAlaGlyValSerTrpGlyLysAlaGln 250  
 QY ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 679 CCCAACCGGCCCTGGCATCTACACCCGTCACCTACTACTTGACTGATCCACCAT 738  
 Db 251 ProAsnArgProGlyIleTyrrhArgValThrTyrrhLeuAspTrpIleHisTyr 270  
 QY ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 271 ValProLysLysPro 275

RESULT 3

C3863

C:Species: Homo sapiens (man)  
 C:Accession: C35863; A35863; A38893  
 R:Proc. Natl. Acad. Sci. U.S.A. 87, 3811-3815, 1990  
 A:Title: Human mast cell trypsinase: multiple cDNAs and genes reveal a multigene serine proteinase family.  
 A:Reference number: A35863; MUID:90251647; PMID:2187193  
 A:Accession: C35863  
 A:Molecule type: mRNA  
 A:Residues: 9-275 <VAN>  
 A:Accession: A35863  
 A:Molecule type: DNA  
 A:Residues: 1-9 <VA2>  
 A:Cross-references: GB:M33494; NID:93927804; PIDN: AAC83172.1; PID:9339977  
 A:Note: the first nine residues of this sequence are inferred from genomic DNA of trypsinase.  
 R:Vanderslice, P.  
 submitted to GenBank, April 1990  
 A:Reference number: A38893  
 A:Molecule type: mRNA  
 A:Residues: 9-131, 'K', 132-275 <VA3>  
 A:Cross-references: GB:M33493; NID:9339984; PIDN: AAA36780.1; PID:9339985  
 C:Superfamily: trypsin; trypsin homology  
 C:Keywords: hydrolase; serine protease; zymogen  
 F:1-275/Domain: signal sequence #status predicted <SIG>  
 F:22-30/Domain: activation peptide #status predicted <ACT>  
 F:31-275/Domain: trypsinase I #status predicted <MAT>  
 F:31-267/Domain: trypsin homology <TR>  
 F:74,121,224/Active site: HIS, Asp, Ser #status predicted

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
1	1.09e-84	275	240	0	5	0	2

Db 131 ValAsnValSerSerHisValHisthValThrLeuProProLaserGluThrPhePro 150  
 QY ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 19 ATCTCGGGGGTAGGAGGCCAGGCCAGGACAAGTGCCCTGGCAGGAGGCAGTC 78  
 Db 31 LeuValGlyGlyLysGlnValAlaProArgSerLeuLysTrpProTrpGlnValSerLeuArgVal 50  
 QY ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 79 CAGGCCATACTGGATCCTGGCTGGGGGCTCCCATCACCCAGGGCTGTG 138  
 Db 51 ArgAsparGlyTrpMetHisPheCysGlyLysSerLeuLeuLeuHisProGlnTrpValIeu 70  
 QY ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 139 ACCCGCCGCCGCGCTGGCAAGGCCGACGTCAGCATCTGCCGCCCCTAGGTTGCAACTG 198  
 Db 71 ThrAlaAlaHisCysValGlyProAspValAspLeuAlaLeuArgValGlnIeu 90  
 QY ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 259 CACAGTCTACACGCCAGCAGGGAAATGCCATCTGGCAGCTGGAGCTGGAGGCC 318  
 Db 111 ProGlnPhenPheTyrrhAlaGlnLeuGlyAlaAspIleAlaLeuLeuGluIleGluIle 130  
 QY ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 319 GTGAGGCTCAGGCCAGTCACACGGTCAACCTGGCCCTCAGAGCTTCCC 378  
 Db 91 ArgGluGlnHisLeuTyrrhGlyLysGlnLeuLeuLeuProValSerArgIleLeuValIeu 110  
 QY ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 131 ValAsnValSerSerHisValHisthValThrLeuProProLaserGluThrPhePro 150  
 QY ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 439 CCSCCATATCCTGTAAGCAGGTAAGSTCCCTAATGAAACACCATTGTTGAGCA 498  
 Db 171 ProProHeProLeuLysGlnValLysValProLeuMetGluAsnHisIleCysAspAla 190  
 QY ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 191 LYSThrHisLeuGlyAlaTyrrhGlyAspAspValAргLeuValArgLysAspMetIeu 210  
 Db 231 LysValAsnGlyThrTrpLeuGlnAlaGlyValSerTrpGlyLysAlaGln 250  
 QY ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 619 AGGGTGAATGGCACCTGGCTGGCATCTACACCCGTCACCTACTTGACTGATCCACCAT 738  
 Db 251 ProAsnArgProGlyIleTyrrhArgValThrTyrrhLeuAspTrpIleHisTyr 270  
 QY ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 271 ValProLysLysPro 275

RESULT 4

A45754

C:Species: Homo sapiens (man)  
 C:Accession: C45754; B37193  
 R:Miller, J.S.; Westin, E.H.; Schwartz, L.B.  
 J. Clin. Invest. 84, 1188-1195, 1989  
 A:Title: Cloning and characterization of complementary DNA for human trypsinase.  
 A:Reference number: A45754; MUID:90009311; PMID:267049  
 A:Accession: A45754  
 A:Molecule type: mRNA  
 A:Residues: 1-274 <ML>  
 A:Cross-references: GB:M30038  
 R:Miller, J.S.; Moxley, G.; Schwartz, L.B.  
 J. Clin. Invest. 86, 864-870, 1990  
 A:Title: Cloning and characterization of a second complementary DNA for human trypsinase.  
 A:Reference number: A37193; MUID:90369005; PMID:2203827  
 A:Accession: B37193  
 A:Molecule type: mRNA

DS-09-598-982-20 (1-771) x C35863 (1-275)

A;Residues: 1-274 <MI2>  
A;Cross-references: GB:M30038  
A;Note: the sequence from Fig. 4 is inconsistent with that from Fig. 2 in having 205-Prd  
C;Superfamily: trypsin; trypsin homology  
C;Keywords: hydrolase; serine proteinase; zymogen  
F;22-30/Domain: activation peptide #status predicted <SIG>  
F;31-24/Product: trypsinase I #status predicted <ACT>  
F;31-266/Domain: trypsin homology <TRV>  
F;74,120,223/Active site: His, Asp, Ser #status predicted

Alignment Scores:  
Pred. No.: 2.34e-77 Length: 274  
Score: 12.2.50 Matches: 220  
Percent Similarity: 92.65% Conservative: 7  
Best Local Similarity: 89.80% Mismatches: 17  
Query Match: 84.53% Indels: 1  
DB: 2 Gaps: 1

US-09-598-982-20 (1-771) x A45754 (1-274)

Qy 19 ATCGTCGGGGTCAGGAGCCCAAGGACAATGCCCTGGCAGTGAGCTGAGCT 78  
Db 31 IleValGlyGlyGlyGlnGluAlaProArgSerIlyTrpProTrgInValserLeuAlaGln 50  
Qy 79 CACGGCCATACTGGTCACTCTGGGGGTCCCTCATCACCCCCAGTGGGCTCG 138  
Db 51 ArgAspArgGlyTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 70  
Qy 139 ACAGGCCGCGGCGTGGTGGACCGGACTCAAGGATTTGGCCGCCCTCAGGGTGCAC 198  
Db 71 ThrAlaAlaHisCysLeuGlyProAspValLeuAspLeuAlaThrLeuArgValAsn--- 89  
Qy 199 CGGGGAGCAGCCTCTACTACAGGACAGCAGCTGGGTCAGCAGGATCATGTCAC 258  
Db 90 SerGlyThrHistLeuIleTyrgInAspGlnLeuLeuLeuProSerArgIleLeuHis 109  
Qy 259 CCACAGTCTACACGCCAGATCGAACGGACATGCCCTCTGGACCTGGAGGCCG 318  
Db 110 ProGlnPhytyrIleIleGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluIlePro 129  
Qy 319 GTCAGGGCTTCAAGCCACGTCACGTCACCTGGCCCTGGCTCAGAGACCTTCCC 378  
Db 130 ValAsnIleSerSerArgValHistValMetLeuProProAlaSerGluThrPhePro 149  
Qy 379 CCGGGGATGCCGCGTCTGGTCACTGGCTGGAAATATGATGACGCCCTCCA 438  
Db 150 ProGlyMetProCysTrpValIleGlyTrpGlyAspAlaAsnAspGluProLeuPro 169  
Qy 439 CCCCATTCCTGAGCAGGTGAAGTCCCATATGARACACACATTGTGGCCA 498  
Db 170 ProPropheProLeuIleGlyValIleValProIleMetGluAsnHistIleCysAspAla 189  
Qy 499 AAATACCCACCTGGGCCCTACAGGGAGACGAGCTGGTCCGTGACCACTG 558  
Db 190 LysThrIleSleGlyIleIleIleArgAspAspMetLeu 209  
Qy 559 TGGCCGCGAACACCCGAGGGACTCTGGGAGCCCTGGTGGTC 618  
Db 210 CysAlaGlyAsnSerGlnArgAspSerCysIlyGlyGlyProLeuValCys 229  
Qy 619 ANGGTGAATGGCACCCTGGCTGAGGGGGCTGAGCTGGGGCTGCTCC 678  
Db 230 LysValAsnGlyIleThrIleGluGlnIleGlyValSerTrpAspGlyIleCysAlaGln 249  
Qy 679 CCCAACCCGCCCTGGCCTACACCCGCTGCACTACTGGACTGACGATGACCACT 738  
Db 250 ProAsnArgProGlyIleIleIleArgValThrArgValThrIleLeuAspTrpIleHisTyr 269  
Qy 739 GRCGCCAAAAGCCG 753  
Db 270 ValProIleSlySpro 274

A;Residues: 1-276 <RE2>  
A;Cross-references: GB:M57625; NID:9200506; PIDN:AAA39987.1; PID:9200507

A;Note:

the authors translated the codon CCG for residue 24 as Ala, GAG for residue 3 as Gly, GAG for residue 148 as Gly, and GAA for residue 168 as Gly, and GAA for 185 as Gly

A;Title: Cloning of the cDNA and gene of mouse mast cell protease-6. Transcription by

A;Reference number: A38654; MUID:91139682; PMID:1995638

A;Accession: A38654

A;Molecule type: DNA

A;Residues: 1-276 <RE2>

A;Cross-references: GB:M57626; NID:9200508; PIDN:AAA39988.1; PID:9200509

R;Reynolds, D.S.; Stevens, R.L.; Lane, W.S.; Carr, M.H.; Austen, K.F.; Serafin, W.E.

Proc. Natl. Acad. Sci. U.S.A. 87, 3230-3234, 1990

A;Title: Different mouse mast cell populations express various combinations of at lea

A;Reference number: A35646; MUID:9022202; PMID:2356280

A;Accession: D35646

A;Molecule type: protein

A;Residues: 32-54 <RE3>

R;Huang, R.; Abrink, M.; Gobl, A.E.; Nilsson, G.; Aveskogh, M.; Larsson, L.G.; Nilss

Scand. J. Immunol. 38, 359-367, 1993

A;Title: Expression of a mast cell tryptase in the human monocytic cell lines U-937 a

A;Reference number: 159478; MUID:94023807; PMID:8210998

A;Accession: 159478

A;Title: Preliminary

A;Molecule type: mRNA

A;Residues: 1-276 <RES>

A;Cross-references: GB:L31853; NID:9473480; PIDN:AAA39725.1; PID:9473481

C;Genetics:

A;Gene: MmcP-6

A;Introns: 24/1; 79/2; 168/1; 222/3

C;Superfamily: trypsin; trypsin homology

C;Keywords: hydrolase; serine proteinase; zymogen

F;22-31/Domain: activation peptide #status predicted <ACT>

F;32-276/Product: mast cell proteinase 6 #status experimental

F;32-268/Domain: trypsin homology <TRV>

F;75, 122, 225/Active site: His, Asp, Ser #status predicted

A;Cross-references:

A;Gene: MmcP-6

A;Introns: 24/1; 79/2; 168/1; 222/3

C;Superfamily: trypsin; trypsin homology

C;Keywords: hydrolase; serine proteinase; zymogen

F;22-31/Domain: activation peptide #status predicted <ACT>

F;32-276/Product: mast cell proteinase 6 #status experimental

F;32-268/Domain: trypsin homology <TRV>

F;75, 122, 225/Active site: His, Asp, Ser #status predicted

A;Cross-references:

A;Gene: MmcP-6

A;Introns: 24/1; 79/2; 168/1; 222/3

C;Superfamily: trypsin; trypsin homology

C;Keywords: hydrolase; serine proteinase; zymogen

F;22-31/Domain: activation peptide #status predicted <ACT>

F;32-276/Product: mast cell proteinase 6 #status experimental

F;32-268/Domain: trypsin homology <TRV>

F;75, 122, 225/Active site: His, Asp, Ser #status predicted

A;Cross-references:

A;Gene: MmcP-6

A;Introns: 24/1; 79/2; 168/1; 222/3

C;Superfamily: trypsin; trypsin homology

C;Keywords: hydrolase; serine proteinase; zymogen

F;22-31/Domain: activation peptide #status predicted <ACT>

F;32-276/Product: mast cell proteinase 6 #status experimental

F;32-268/Domain: trypsin homology <TRV>

F;75, 122, 225/Active site: His, Asp, Ser #status predicted

A;Cross-references:

A;Gene: MmcP-6

A;Introns: 24/1; 79/2; 168/1; 222/3

C;Superfamily: trypsin; trypsin homology

C;Keywords: hydrolase; serine proteinase; zymogen

F;22-31/Domain: activation peptide #status predicted <ACT>

F;32-276/Product: mast cell proteinase 6 #status experimental

F;32-268/Domain: trypsin homology <TRV>

F;75, 122, 225/Active site: His, Asp, Ser #status predicted

A;Cross-references:

A;Gene: MmcP-6

A;Introns: 24/1; 79/2; 168/1; 222/3

C;Superfamily: trypsin; trypsin homology

C;Keywords: hydrolase; serine proteinase; zymogen

F;22-31/Domain: activation peptide #status predicted <ACT>

F;32-276/Product: mast cell proteinase 6 #status experimental

F;32-268/Domain: trypsin homology <TRV>

F;75, 122, 225/Active site: His, Asp, Ser #status predicted

A;Cross-references:

A;Gene: MmcP-6

A;Introns: 24/1; 79/2; 168/1; 222/3

C;Superfamily: trypsin; trypsin homology

C;Keywords: hydrolase; serine proteinase; zymogen

F;22-31/Domain: activation peptide #status predicted <ACT>

F;32-276/Product: mast cell proteinase 6 #status experimental

F;32-268/Domain: trypsin homology <TRV>

F;75, 122, 225/Active site: His, Asp, Ser #status predicted

A;Cross-references:

A;Gene: MmcP-6

A;Introns: 24/1; 79/2; 168/1; 222/3

C;Superfamily: trypsin; trypsin homology

C;Keywords: hydrolase; serine proteinase; zymogen

F;22-31/Domain: activation peptide #status predicted <ACT>

F;32-276/Product: mast cell proteinase 6 #status experimental

F;32-268/Domain: trypsin homology <TRV>

F;75, 122, 225/Active site: His, Asp, Ser #status predicted

A;Cross-references:

A;Gene: MmcP-6

A;Introns: 24/1; 79/2; 168/1; 222/3

C;Superfamily: trypsin; trypsin homology

C;Keywords: hydrolase; serine proteinase; zymogen

F;22-31/Domain: activation peptide #status predicted <ACT>

F;32-276/Product: mast cell proteinase 6 #status experimental

F;32-268/Domain: trypsin homology <TRV>

F;75, 122, 225/Active site: His, Asp, Ser #status predicted

A;Cross-references:

A;Gene: MmcP-6

A;Introns: 24/1; 79/2; 168/1; 222/3

C;Superfamily: trypsin; trypsin homology

C;Keywords: hydrolase; serine proteinase; zymogen

F;22-31/Domain: activation peptide #status predicted <ACT>

F;32-276/Product: mast cell proteinase 6 #status experimental

F;32-268/Domain: trypsin homology <TRV>

F;75, 122, 225/Active site: His, Asp, Ser #status predicted

A;Cross-references:

A;Gene: MmcP-6

A;Introns: 24/1; 79/2; 168/1; 222/3

C;Superfamily: trypsin; trypsin homology

C;Keywords: hydrolase; serine proteinase; zymogen

F;22-31/Domain: activation peptide #status predicted <ACT>

F;32-276/Product: mast cell proteinase 6 #status experimental

F;32-268/Domain: trypsin homology <TRV>

F;75, 122, 225/Active site: His, Asp, Ser #status predicted

A;Cross-references:

A;Gene: MmcP-6

A;Introns: 24/1; 79/2; 168/1; 222/3

C;Superfamily: trypsin; trypsin homology

C;Keywords: hydrolase; serine proteinase; zymogen

F;22-31/Domain: activation peptide #status predicted <ACT>

F;32-276/Product: mast cell proteinase 6 #status experimental

F;32-268/Domain: trypsin homology <TRV>

F;75, 122, 225/Active site: His, Asp, Ser #status predicted

A;Cross-references:

A;Gene: MmcP-6

A;Introns: 24/1; 79/2; 168/1; 222/3

C;Superfamily: trypsin; trypsin homology

C;Keywords: hydrolase; serine proteinase; zymogen

F;22-31/Domain: activation peptide #status predicted <ACT>

F;32-276/Product: mast cell proteinase 6 #status experimental

F;32-268/Domain: trypsin homology <TRV>

F;75, 122, 225/Active site: His, Asp, Ser #status predicted

A;Cross-references:

A;Gene: MmcP-6

A;Introns: 24/1; 79/2; 168/1; 222/3

C;Superfamily: trypsin; trypsin homology

C;Keywords: hydrolase; serine proteinase; zymogen

F;22-31/Domain: activation peptide #status predicted <ACT>

F;32-276/Product: mast cell proteinase 6 #status experimental

F;32-268/Domain: trypsin homology <TRV>

F;75, 122, 225/Active site: His, Asp, Ser #status predicted

A;Cross-references:

A;Gene: MmcP-6

A;Introns: 24/1; 79/2; 168/1; 222/3

C;Superfamily: trypsin; trypsin homology

C;Keywords: hydrolase; serine proteinase; zymogen

F;22-31/Domain: activation peptide #status predicted <ACT>

F;32-276/Product: mast cell proteinase 6 #status experimental

F;32-268/Domain: trypsin homology <TRV>

F;75, 122, 225/Active site: His, Asp, Ser #status predicted

A;Cross-references:

A;Gene: MmcP-6

A;Introns: 24/1; 79/2; 168/1; 222/3

C;Superfamily: trypsin; trypsin homology

C;Keywords: hydrolase; serine proteinase; zymogen

F;22-31/Domain: activation peptide #status predicted <ACT>

F;32-276/Product: mast cell proteinase 6 #status experimental

F;32-268/Domain: trypsin homology <TRV>

F;75, 122, 225/Active site: His, Asp, Ser #status predicted

A;Cross-references:

A;Gene: MmcP-6

A;Introns: 24/1; 79/2; 168/1; 222/3

C;Superfamily: trypsin; trypsin homology

C;Keywords: hydrolase; serine proteinase; zymogen

F;22-31/Domain: activation peptide #status predicted <ACT>

F;32-276/Product: mast cell proteinase 6 #status experimental

F;32-268/Domain: trypsin homology <TRV>

F;75, 122, 225/Active site: His, Asp, Ser #status predicted

A;Cross-references:

A;Gene: MmcP-6

A;Introns: 24/1; 79/2; 168/1; 222/3

C;Superfamily: trypsin; trypsin homology

C;Keywords: hydrolase; serine proteinase; zymogen

F;22-31/Domain: activation peptide #status predicted <ACT>

F;32-276/Product: mast cell proteinase 6 #status experimental





A; Reference number: JC4171; MUID:96015171; PMID:8537314  
A; Accession: JC4171  
A; Molecule type: mRNA  
A; Residues: 1-274 <IDE>  
A; Cross-references: DDBJ:D38455; NID:9556555; PIDN:BAA07486.1; PID:9556556  
C; Comment: This enzyme is basically specific for a connective tissue mast cell, it is up-einase inhibitors.  
C; Superfamily: trypsin; trypsin homology  
C; Keywords: glycoprotein; hydrolase; mast cell; serine proteinase; zymogen  
F; 1-19/Domain: signal sequence #status predicted <SIG>  
F; 20-29/Domain: activation peptide #status predicted <ACT>  
F; 30-266/Domain: mast cell trypsinase #status predicted <MAT>  
F; 73,120,223/Active site: His, Asp, Ser #status predicted  
F; 131/Binding site: carbohydrate (Asn) (covalent) #status predicted  
Alignment Scores:  
Pred. No.: 3.92e-65 Length: 274  
Score: 1055.00 Matches: 187  
Percent Similarity: 84.06% Conservative: 24  
Best Local Similarity: 74.50% Mismatches: 38  
Query Match: 72.36% Indels: 2  
DB: 2 Gaps: 1

US-09-598-982-20 (1-771) x JC4171 (1-274)

Qy 4 CCCCTCGAGAAAAGA-----ATGTCGGGGTCAGGAGCCCCAGGGCACTGGCC 57  
||| :::::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 23 ProvalysGlnArgValGlyIleValGlyGlyArgGluAlaSerGluLeuTrpPro 42

Qy 58 TGCAGGTGAGCTGAGAGTCCACGGCCCATACTGGATGACTCTGGGGGCCCC 117  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 43 TrpGlnValSerLeuArgPhelysPheSerPheTrpMetHisPheCysGlyGlySerIle 62

Qy 118 ATTCACCCCACTGGGCTGTACGCCCGGGCTGCGGGACGGACGTCAGGATCG 177  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 63 IleHistGlnIlePheLeuAlanLysValGlyLeuHistIleLysSerPro 82

Qy 178 GCGCCCTCAGGGTCAACTGGGGAGCACCTCTACCTACGGACAGCTGGCC 237  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 83 GluteuPheArgValGlnIleLeuArgGluInTyLeuTyTyRalaAspGlnIleLeuThr 102

Qy 238 GTCAGCAGGATCATGCCAACACAGTCTACAGCCGCCAGATGGACGCCATGCC 297  
||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 103 ValAsnArgThrValValHisProHISProHISProHISProHISProHIS 122

Qy 298 CAGCTGGAGCTGGAGGACCGCGGAAGCTCTCCAGCCAGTCCACACGCTCACCTGCC 357  
||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 123 LeuLeuGluLeuGluIleProValAsnValSerIleSerThrHisIleHisProLeuPro 142

Qy 358 CCTGCCCTCAGAGACTTCCCCGGGGATGCCCTGCTACTGGCTACTGGCTGATG 417  
||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 143 ProLaserGluLthrPheProSerGlyLhrSerCystTrpValThrGlyTrpGlyAspIle 162

Qy 418 GACAATGATGAGGGCTCCACGCCATTCTCTGAAACGAGGTGGAGGCCATAATG 477  
||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 163 AspSerAspGluProLeuLeuProLeuProLeuProLeuProLeuPro 182

Qy 478 GAAACACATATGTGACGCAAATACACCTTGGCCGCTACAGGGAGACGAGTCCC 537  
||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 183 GluAsnSerLeuCysAspArgLysTyrIleIleLeuTyrGlyAspAspValPro 202

Qy 538 ATGTCCTGAGCACATCTGCTGCTGGGGAGACACCCGGAGACTCATGCCAGGGCAC 597  
||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 203 IleLeuGlnAspGlyMetLeuCysAlaGlyAsnThrAspSerCysGlyLysAsp 222

Qy 598 TCCGAGGGCCCTGGTGTCGAAGGTGAATGGCACCTTGCTGCGGCCGCGTGTAC 657  
||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 223 SerGlyGlyProLeuValCysLysValLysGlyThrIleLeuGluAlaGlyValSer 242

Qy 658 TGGGCCAGGCGCTGCCAGGCCAACCGCCGCGCTACACCCGCGTGTAC 717  
||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 243 TrpGlyGluGlyCysAlaGluAlaAsnGlyProGlyIleTyrThrProValThrTyr 262

Qy 718 TGGACGCTGAGTCCACCATATGCCAAAAAG 750  
||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 263 LieAspAspPheIleHisArgTyrValProGlnArg 273

RESULT 10

S68702 tryptase (EC 3.4.21.59) - bovine (fragment)  
C; Species: Bos primigenius taurus (cattle)  
C; Date: 25-Feb-1998 #sequence\_revision 13-Mar-1998 #text\_change 20-Jun-2000  
C; Accession: S68702; S68676  
R; Pallaor, M.; Gambacorta, A.; Fiorucci, L.; Mignogna, G.; Barra, D.; Ascoli, F.  
Eur. J. Biochem. 237, 10-15, 1996  
A; Title: cDNA cloning and primary structure of trypase from bovine mast cells, and  
A; Reference number: S68676; MUID:96203914; PMID:8620861  
A; Accession: S68702  
A; Molecule type: mRNA  
A; Residues: 1-237 <PAI>  
A; Cross-references: EMBL:X94982; NID:91332446; PIDN:CAA6438.1; PID:91332447  
A; Molecule type: protein  
A; Residues: 1VSGQEAP,1-8;61-74;9-97;126-148;162-190;208-222;227-235 <PAW>  
C; Superfamily: trypsin; trypsin homology  
C; Keywords: glycoprotein; hydrolase; serine proteinase; zymogen  
F; 1-228/Domain: trypsin; trypsin homology (fragment); <TR>  
F; 36,83,177/Active site: His, Asp, Ser #status predicted  
F; 94,195/Binding site: carbohydrate (Asn) (covalent) #status predicted  
Alignment Scores:  
Pred. No.: 1.09e-61 Length: 237  
Score: 1005.00 Matches: 178  
Percent Similarity: 84.32% Conservative: 21  
Best Local Similarity: 75.42% Mismatches: 37  
Query Match: 68.93% Indels: 0  
DB: 2 Gaps: 0

US-09-598-982-20 (1-771) x S68702 (1-237)

Qy 46 ACCAGTGGCCCTGAGGTGAGAGTCCACGGCCCATACTGGATGACTCTGC 105  
||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 2 SerGlnTrpProTrpGlnValSerIleLeuArgValSerArgArgTyrTrpArgHisCys 21

Qy 106 GGGGCTCTCATCCGCCAGTGGTGTGCAACCCGCCGGGGTGGGGACGCCGAC 165  
||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 22 GlyGlySerLeuIleHisProGlnTrpValLeuThrAlaAlaLysValGlyProGlu 41

Qy 166 GTCAGGATCTGCCGCGCTAGGGTCAACTGGGGAGACGCCCTACTACAGGC 225  
||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 42 ValHisGlyProSerTrpHeArgValGlnIleArgGluGlyIleAspValPro 61

Qy 226 CAGCTGCTGCCGTCAGGAGTCATGTGACCCAGTCAGTCTACACGCCAGATCGA 285  
||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 62 GluLeuLeuProIleSerArgIleLeuProHisProAsnCysTyrSerValLysAspGly 81

Qy 286 GGGGACATCGCCCTGGCGAGGATGGAGGCCGGTGTAGGTCCTCCAGCCACGCCAG 345  
||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 82 AlaAspIleAlaLeuLeuGluLeuAspLysLeuValAlaAsnIleSerTrpHisValGlnPro 101

Qy 346 GTCACCTGCCCCCTGGCTAGAGACCTTCCCCGGGATGCCGTGTGGGCCACTGGC 405  
||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 102 ValIleLeuProProGluSerGluIleThrPheProProGlyThrGlnCysTrpValThrGly 121

Qy 406 TGGGCATGTTGACATGATGAGGCCCTCCACGCCATTCTCTGAGCAGGTGAAG 465  
||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 122 TrpGlyIleAspAspGlyLysArgGluLeuProProPheProIleLeuAspValIle 141

Qy 466 GTCCTCATATACTGGAAACCACTTGTGACGCAAATACACCTTGGGCCACTACCGGA 525  
||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 142 ValProValValGluAsnSerValCysAspArgLysTyrIleSerGlyLeuSerThrGly 161

Qy 526 GACGACTCGCATGCCGTCGACGACATGGCTGCTGGGGACACCCGGAGGACTCA 585  
||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 162 AspAspIleProIleValArgLysAspMetLeuCysAlaGlyAspSerGlyAspIle 181

QY 586 TGCAGGGCCGACTCGGGAGGGCCCCCTGGTGCAGGAAATGGCACCTGGCTGAGGCC 645  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 182 CysGlyLysAspSerGlyGlyProLeuValAsnGlyThrPheUglnala 201  
 QY 546 GCGGTGGCAGGTGGCCGAGGGCTGCCCCAACGGCTGSGCATCACACCGT 705  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 202 GlyValValSerGlyGlyProLysAlaLysProAsnArgProGlyIleTyrThrArg 221  
 QY 706 GTCACCTACTTGACTGGATCCACACTATGCCCAAAAGCG 753  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 222 ValThrSerTyrLeuAspTripleHsGlnTyrValProGlyPro 237  
 RESULT 11  
 I48685 mast cell proteinase 6 (EC 3.4.21.-) precursor - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Accession: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 22-Jun-1999  
 C:Accession: I48685; S43172  
 Immunogenetics for mast-cell serine protease and their molecular evolution.  
 A:Title: Genes for mast-cell serine protease and their molecular evolution.  
 A:Reference number: 148684; MUID:95048582; PMID:7959952  
 A:Status: preliminary; translated from GB/EMBL/DDJB  
 A:Molecule type: mRNA  
 A:Residues: 40 397-414, 1994  
 A:Cross-references: EMBL:X78542.; NID:9468809; PIDN:CAA55288.1; PID:9468810  
 C:Superfamily: trypsin; trypsin homology  
 F:32-230/Domain: trypsin homology #status atypical <TRY>  
 Alignment Scores:  
 Pred. No.: 4 03e-48 Length: 230  
 Score: 808.00 Matches: 144  
 Percent Similarity: 81.82% Conservative: 18  
 Best Local Similarity: 72.73% Mismatches: 34  
 Query Match: 55.42% Indels: 2  
 DB: 2 Gaps: 1  
 US-09-598-982-20 (1-771) x I48685 (1-230)  
 QY 4 CCGCTCGAGAAAAGA-----ATCGTGGGGCTGAGGAGGAGGAGAGTGCCC 57  
 ||||| ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 25 ProAlaAsnGlnArgValGlyIleValGlyIlyHisGluSerGluSerIlyTrpPro 44  
 QY 58 TGGCAGGTGGCTGAGTCCACGCCATCTGGTGAACACTCCCGGGCTCCCT 117  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 45 TrpGlyIvalSerLeuArgPhelysLeuAsnTyrrTripleHsPheDysGlyIlySerIle 64  
 QY 19 ATCGCTCGGGGTCAGGAGGCCAGAGCAAGTGGCCCTGGCAGSTGAGAGTC 78  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 20 IleValGlyGlyCysLysValProAlaArgArgTyrProTrpGlnValSerLeuArgPhe 39  
 QY 79 CGGG-----CCATACTGGTGCACCTCTGGGGCCCTCCATCCCCCG 129  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 40 HisGlyMetGlySerGlyGlnTrpGlnHistidylSerGlySerLeuIleHisProGln 59  
 QY 130 TGGGTGCTGACGCCGGGGTGGTGGACGGCTCAAGGATCTGGCCCTCAGG 189  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 60 TrpValIleThrAlaAlaHisCysValGluIleGlyLeuGluAlaAlaThrLeuArg 79  
 QY 190 GRCGAACCTGCGGAGCAGCACCTCTACTACAGGACAGCTGCTGCGATGAGATC 249  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 80 ValGlnValGlyGlnIleArgLeuTyrAspHisAspGlnIleCysAsnValThrGluIle 99  
 QY 250 ATGGTGCACCCACAGTCTACCCGCCAGATCGGA-----CGGACATGCCCTG 300  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 100 IleArgIlysProAsnProAsnMetSerTyrPheAsnMetSerTyrPheAsnMetSerTyrPheAsn 119  
 Db 105 LeuAsnArgIleLeuValIleHisProIleValIleTyrIleGlyIlyAlaAspValAla 124  
 QY 358 CTGGCTCAGGACGACCTTCCCGGGATGCGTCACTGCTGGGGCATGGTGC 417  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 298 CTCGCTGAGCTGGAGGAGGCCGAGCTCCAGGACAGTCACCGTCACCC 357  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 125 LeuLeuGluLeuValProValAsnValSerThrHistidylSerIleSerLeuPro 144  
 QY 418 GACAATGATGAGGCCCTCCACGCCATTTCTCTGAGCAGGTGAAGGTGCCCATATG 477  
 ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 165 AspAsnAspGluProLeuProProProtYProLeuWysGlnValLysValProIleVal 184  
 QY 478 GAAACCACATTGAGCACAATACACCTTGGGCCCTACACGGAGACGACGCC 537  
 ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 185 GluAsnSerLeuCysAspArgLysTyrHistidylGlyLeuTyrThrGlyAspAspPhePro 204  
 QY 538 ATGTCGGTGAAGGACATGCTGCGGGACACGGAGGACTGATGCCAG 591  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 205 IleValHisAspGlyMetLeuCysAlaGlyAsnThrArgArgAspSerCysGln 222  
 RESULT 12  
 B32410 mastcytoma proteinase (EC 3.4.21.-) precursor - dog  
 C:Species: Canis lupus familiaris (dog)  
 C:Accession: 12-Oct-1989 #sequence\_revision 31-Dec-1993 #text\_change 10-Sep-1997  
 R:Vanderline, P.; Craik, C.S.; Nadel, J.A.; Caughey, G.H.  
 Biochemistry 28, 4148-4155, 1989  
 A:Title: Molecular cloning of dog mast cell trypase and a related protease: structure  
 A:Reference number: A32410; MUID:89352460; PMID:2504277  
 A:Accession: B32410  
 A:Molecule type: mRNA  
 A:Cross-references: GB:M24665; NID:9163984; PID:9163985; GB:J02862  
 A:Note: the authors translated the codon AGC for residue 114 as Ser  
 C:Superfamily: trypsin; trypsin homology  
 C:Keywords: hydrolase; serine proteinase  
 F:1-19/Domain: signal sequence #status predicted <SGN>  
 F:20-269/Product: mastcytoma proteinase #status predicted <MAT>  
 F:20-259/Domain: trypsin homology <TRY>  
 F:66,116,217/Active site: His, Asp, Ser #status predicted  
 Alignment Scores:  
 Pred. No.: 1 2.28e-39 Length: 269  
 Score: 684.50 Matches: 132  
 Percent Similarity: 65.34% Conservative: 32  
 Best Local Similarity: 52.59% Mismatches: 78  
 Query Match: 46.95% Indels: 9  
 DB: 2 Gaps: 4  
 US-09-598-982-20 (1-771) x B32410 (1-269)  
 QY 19 ATCGCTCGGGGTCAGGAGGCCAGAGCAAGTGGCCCTGGCAGSTGAGAGTC 78  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 20 IleValGlyGlyCysLysValProAlaArgArgTyrProTrpGlnValSerLeuArgPhe 39  
 QY 79 CGGG-----CCATACTGGTGCACCTCTGGGGCCCTCCATCCCCCG 129  
 ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 40 HisGlyMetGlySerGlyGlnTrpGlnHistidylSerGlySerLeuIleHisProGln 59  
 QY 130 TGGGTGCTGACGCCGGGGTGGTGGACGGCTCAAGGATCTGGCCCTCAGG 189  
 ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 60 TrpValIleThrAlaAlaHisCysValGluIleGlyLeuGluAlaAlaThrLeuArg 79  
 QY 190 GRCGAACCTGCGGAGCAGCACCTCTACTACAGGACAGCTGCTGCGATGAGATC 249  
 ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 80 ValGlnValGlyGlnIleArgLeuTyrAspHisAspGlnIleCysAsnValThrGluIle 99  
 QY 250 ATGGTGCACCCACAGTCTACCCGCCAGATCGGA-----CGGACATGCCCTG 300  
 ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 100 IleArgIlysProAsnProAsnMetSerTyrPheAsnMetSerTyrPheAsnMetSerTyrPheAsn 119  
 QY 301 CTGGAGCTGGAGGAGCGGCTGGAGGCTCAGGCCAGTCACCGTCACCC 360  
 ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 120 LeuLysIleGluAlaProLeuThrLeuSerGluAspValAsnLeuValSerLeuProSer 139  
 QY 361 GCTCAAGAGACCTTCCCGGGATGCGTCACTGCTGGGGCATGGTGC 420  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 140 ProSerLeuIleValProProGlyMetLeuCysTrpValIleGlyIlyAspIleAla 159  
 QY 421 ATGAGAGGCCCTCCACGCCATTTCTCTGAGCAGGTGAAGGTGCCCATATG 480  
 ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 160 AspHisIhrProLeuProProProtYProLeuWysGlnValGlyAlaProIleVal 179



R McMullen, B.A.; Fujikawa, K.; Davie, E.W.  
 Biochemistry 30, 2056-2060, 1991  
 A;Title: Location of the disulfide bonds in human coagulation factor XI: the presence of  
 A;Reference number: A37940; MUID:91152017; PMID:1998667  
 A;Accession: A37940  
 A;Molecule type: protein  
 A;Residues: '28-'33-'35-'49, 'X' '51-'55, 'X' '57-'63-'70-'75, 'X' '77-'79-'107-'109, 'X' '111-'112-'132-'139;  
 '280-'282, 'X' '284-'285-'297, '313-'316, 'X' '318-'319-'320-'326, 'X' '328-'330, 'X' '347-'349-'373, 'X' '375  
 C;Comment: The proenzyme consists of two identical chains linked by one or more disulfide bonds at the active site, and a heavy chain, which associates with high molecular weight (HMW) kinins.  
 C;Genetics: A;Gene: GDB:FI1  
 A;Cross-references: GDB:119891; OMIM:264900  
 A;Map position: 4q35-4q35  
 A;Introns: 19/1; 21/2; 109/1; 162/2; 199/1; 252/2; 289/1; 343/2; 379/1; 435/2; 494/1; 52  
 C;Function:  
 A;Description: catalyzes the proteolytic activation of coagulation factor IX  
 A;Pathway: blood coagulation intrinsic pathway  
 C;superfamily: coagulation factor XI: trypsin homology  
 C;Keywords: blood coagulation; duplication; glycoprotein; hemophilia C; homodimer; hydrolysis; signal sequence  
 F;18-/Domain: signal sequence #status predicted <SG>  
 F;19-387/Product: coagulation factor XIA heavy chain #status experimental <HCH>  
 F;19-108/Domain: apple repeat <API>  
 F;19-109-198/Domain: apple repeat <API>  
 F;199-288/Domain: apple repeat <API>  
 F;290-379/Domain: apple repeat <API>  
 F;388-625/Product: coagulation factor Xia light chain #status experimental <LCH>  
 F;388-618/Domain: trypsin homology <TR>  
 F;29-Disulfide bonds: interchain #status experimental  
 F;46-76,50-56,110-193,136-165,140-146,200-283,226-255,230-236,291-374,317-346,321-327,38  
 F;46-76,53,45,46,50-59/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F;39-/Disulfide bonds: interchain #status predicted  
 F;387-388/Cleavage site: Arg-Tle (coagulation factor XIA) #status experimental  
 F;431-480,575/Active site: His, Asp, Ser #status predicted  
 F;491/Binding site: carbohydrate (Asn) (covalent) #status experimental  
 Alignment Scores:  
 Pred. No.: 1.18e-24 Length: 625  
 Score: 467.00 Matches: 99  
 Percent Similarity: 51.59% Conservative: 31  
 Best Local Similarity: 39.29% Mismatches: 94  
 Query Match: 32.03% Indels: 28  
 DB: 1 Gaps: 6 1 US-09-598-982-20 (1-771) x KFHU1 (1-625)  
 Result 15  
 KOHP  
 plasma kallikrein (EC 3.4.21.34) precursor - human  
 N;Alternate names: kininogenin; plasma prekallikrein  
 C;Species: Homo sapiens (man)  
 C;Date: 13-Aug-1986 #sequence\_revision 13-Aug-1986 #text\_change 18-Jun-1999  
 C;Accession: A00921; A37939  
 R;Chung, D.W.; Fujikawa, K.; McMullen, B.A.; Davie, E.W.  
 Biochemistry 25, 2410-2417, 1986  
 A;Title: Human plasma prekallikrein, a zymogen to a serine protease that contains four disulfide bonds  
 A;Reference number: A00921; MUID:86243359; PMID:3521732  
 A;Accession: A00921  
 A;Molecule type: mRNA  
 A;Residues: 1-638 <CHU>  
 A;Cross references: GM:13143; NID:gi190262; PIDN:AAA60153.1; PID:gi190263  
 R;McMullen, B.A.; Fujikawa, K.; Davie, E.W.  
 Biochemistry 30, 2050-2056, 1991  
 A;Title: Location of the disulfide bonds in human plasma prekallikrein: the presence of four disulfide bonds  
 A;Reference number: A37939; MUID:91152016; PMID:1998666  
 A;Accession: A37939  
 A;Molecule type: protein  
 A;Residues: '20-'27-'40-'46, 'X' '48, 'H' '50, 'X' '52-'70, 'H' '75-'76, 'X' '78-'80; 103-'113-'131-'140; 1  
 '260-'283, 'X' '285-'287-'291, 'X' '293-'295-'314-'317, 'X' '318-'320-'321-'324, 'X' '329-'333-'334-'339,  
 '55-'58-'59-'62, 'X' '564-'567-'573, 'X' '575-'576-'585-'592-'604 <MDP>  
 C;Comment: This protein, synthesized in the liver, circulates as a noncovalent complex and is activated by factor XIIa, which cleaves the molecule into a zymogen and may also play a role in the renin-angiotensin system by converting prorenin to renin.  
 C;Genetics: A;Gene: GDB:K13  
 A;Cross References: GDB:127575; OMIM:229000  
 A;Map position: 4q35-4q35  
 C;Superfamily: coagulation factor XI; trypsin homology  
 F;1-19/Domain: blood coagulation; duplication; fibrinolysis; glycoprotein; hydrolase; in  
 A;Map position: 4q35-4q35  
 C;Keywords: blood coagulation; signal sequence #status predicted <SG>  
 F;20-38/Domain: plasma kallikrein #status predicted <MAT>  
 F;20-39/Domain: plasma kallikrein heavy chain #status predicted <HCH>  
 F;20-109/Domain: apple repeat <API>  
 F;110-199/Domain: apple repeat <API>  
 F;200-289/Domain: apple repeat <API>  
 F;391-638/Domain: plasma kallikrein light chain #status predicted <LCH>  
 F;391-621/Domain: trypsin homology <TR>  
 F;391-104,47-'57-'111-'194,131-'166-'141-'147,201-'284,221-'256,231-'337,292-'375,322-'328,  
 F;327,308,396,453-'494/Binding site: carbohydrate (Asn) (covalent) #status experimental  
 F;318-'347,340-'345/Disulfide bonds: #status predicted

F:390-391/Cleavage site: Arg-Ile (coagulation factor XIIa) #status predicted  
F:434 483 578/Active site: His, Asp, Ser #status predicted

Alignment:	Scores:
Pred. NO.:	1
Score:	4.9e-23
Percent Similarity:	451.00
Best Local Similarity:	54.40%
Query Match:	38.40%
DB:	30.93%
Length:	638
Matches:	96
Conservative:	40
Mismatches:	84
Indels:	30
Gaps:	8

03-03-338-582-20 (I-1/I) X KUHOF (I-838)

Search completed: January 31, 2003, 06:59:25  
Job time : 33.5 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: January 31, 2003, 05:19:47 ; Search time 62 Seconds

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Perfect score: 1.99gccttcgagaaaaaat.....cgtgaagccggccgcgt 771

Sequence:

Scoring table: IDENTITY\_NUC Gapop 10.0 , Gapext 1.0

Searched: 396772 seqs, 224632407 residues

Total number of hits satisfying chosen parameters: 793544

Minimum DB seq length: 0 Maximum DB seq length: 0

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:\*

```

1: /cgn2_5/podata/2/pubpna/us07_pubcomb.seq:*
2: /cgn2_6/podata/2/pubpna/pctc_new_pub.seq:*
3: /cgn2_6/podata/2/pubpna/us06_new_pub.seq:*
4: /cgn2_6/podata/2/pubpna/us05_pubcomb.seq:*
5: /cgn2_6/podata/2/pubpna/us01_new_pub.seq:*
6: /cgn2_6/podata/2/pubpna/pctus_pubcomb.seq:*
7: /cgn2_6/podata/2/pubpna/us08_new_pub.seq:*
8: /cgn2_6/podata/2/pubpna/us08_pubcomb.seq:*
9: /cgn2_6/podata/2/pubpna/us09_new_pub.seq:*
10: /cgn2_6/podata/2/pubpna/us09_pubcomb.seq:*
11: /cgn2_6/podata/2/pubpna/us10_new_pub.seq:*
12: /cgn2_6/podata/2/pubpna/us10_pubcomb.seq:*
13: /cgn2_6/podata/2/pubpna/us10_pubseq:*
14: /cgn2_6/podata/2/pubpna/us60_pubcomb.seq:*

```

Pred No is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
c	1	726.2	94.2	1081 10 US-09-954-456-2126 Sequence 2126, Appl
c	2	223.2	28.9	616 10 US-09-954-456-2242 Sequence 2242, Appl
c	3	201.4	26.1	1110 9 US-10-041-006A-1 Sequence 1, Appl
c	4	201.4	26.1	1110 12 US-10-040-655-1 Sequence 1, Appl
c	5	201.4	26.1	1129 9 US-10-028-072-221 Sequence 221, Appl
c	6	193.8	25.1	1130 9 US-10-041-006A-8 Sequence 8, Appl
c	7	193.8	25.1	1130 12 US-10-040-655-8 Sequence 8, Appl
c	8	191.6	24.9	2847 10 US-09-888-615-35 Sequence 35, Appl
c	9	187.8	24.4	1613 12 US-10-041-400A-1 Sequence 1, Appl
c	10	187.8	24.4	1613 12 US-10-041-264A-1 Sequence 1, Appl
c	11	187.8	24.4	1613 12 US-10-042-091A-1 Sequence 1, Appl
c	12	187.6	24.3	1122 10 US-09-900-754-1 Sequence 1, Appl
c	13	184.4	23.9	249 10 US-09-960-352-1410 Sequence 1410, Appl
c	14	181.2	23.5	1130 12 US-10-041-400A-8 Sequence 8, Appl
c	15	181.2	23.5	1130 12 US-10-041-264A-8 Sequence 8, Appl
c	16	181.2	23.5	1130 12 US-10-042-091A-8 Sequence 9, Appl
c	17	180.4	23.4	714 9 US-10-067-761-9 Sequence 9, Appl
c	18	180.4	23.4	714 10 US-09-804-156-9 Sequence 9, Appl
c	19	180.4	23.3	262 SEQ ID NO: 2126 LENGTH: 1081

RESULT 1	Sequence 11859, A
; Sequence 2126, Application US/09954456	Sequence 56, Appl
; Patent No. US20020115057A1	Sequence 1, Appl
; GENERAL INFORMATION:	Sequence 262, App
; APPLICANT: Young, Paul	Sequence 262, App
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C	Sequence 262, App
; FILE REFERENCE: 689250-76	Sequence 262, App
; CURRENT APPLICATION NUMBER: US/09/954,456	Sequence 262, App
; CURRENT FILING DATE: 2001-09-18	Sequence 1, Appl
; PRIOR APPLICATION NUMBER: US/60/233,617	Sequence 63, Appl
; PRIOR FILING DATE: 2000-01-18	Sequence 2, Appl
; PRIOR APPLICATION NUMBER: US/60/234,052	Sequence 3B, Appl
; PRIOR APPLICATION NUMBER: US/60/234,923	Sequence 2, Appl
; PRIOR FILING DATE: 2000-03-25	Sequence 19, Appl
; PRIOR APPLICATION NUMBER: US/60/235,134	Sequence 256, App
; PRIOR FILING DATE: 2000-03-25	Sequence 256, App
; PRIOR APPLICATION NUMBER: US/60/235,638	Sequence 256, App
; PRIOR FILING DATE: 2000-09-26	Sequence 256, App
; PRIOR APPLICATION NUMBER: US/60/235,711	Sequence 256, App
; PRIOR FILING DATE: 2000-03-27	Sequence 256, App
; PRIOR APPLICATION NUMBER: US/60/235,720	Sequence 256, App
; PRIOR FILING DATE: 2000-09-27	Sequence 256, App
; PRIOR APPLICATION NUMBER: US/60/235,840	Sequence 256, App
; PRIOR FILING DATE: 2000-03-27	Sequence 256, App
; PRIOR APPLICATION NUMBER: US/60/235,863	Sequence 256, App
; PRIOR FILING DATE: 2000-03-27	Sequence 256, App
; NUMBER OF SEQ ID NOS: 2276	Sequence 256, App
; SOFTWARE: PatentIn version 3.0	Sequence 256, App
; SEQ ID NO: 2126	Sequence 256, App
; LENGTH: 1081	Sequence 256, App
; TYPE: DNA	Sequence 256, App
; ORGANISM: Homo sapiens	Sequence 256, App
; US-09-954-456-2126	Sequence 256, App
Query Match Similarity 94.2%; Score 726.2; DB 10; Length 1081;	
Best Local Similarity 98.3%; Pred. No. 7.4e-162; Matches 734; Conservative 0; Mismatches 13; Indels 0; Gaps 0;	

PRIOR APPLICATION NUMBER: US/60/235, 637  
 PRIOR FILING DATE: 2000-09-26  
 PRIOR APPLICATION NUMBER: US/60/235, 638  
 PRIOR FILING DATE: 2000-09-26  
 PRIOR APPLICATION NUMBER: US/60/235, 711  
 PRIOR FILING DATE: 2000-09-27  
 PRIOR APPLICATION NUMBER: US/60/235, 720  
 PRIOR FILING DATE: 2000-09-27  
 PRIOR APPLICATION NUMBER: US/60/235, 840  
 PRIOR FILING DATE: 2000-09-27  
 PRIOR APPLICATION NUMBER: US/60/235, 863  
 PRIOR FILING DATE: 2000-09-27  
 NUMBER OF SEQ ID NOS: 2276  
 SOFTWARE: Patentin version 3.0  
 SEQ ID NO: 2242  
 LENGTH: 616  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE: misc\_feature  
 NAME/KEY: misc\_feature  
 OTHER INFORMATION: n\_a,t\_g or c  
 ; US-09-594-456-2242  
 Query Match 28.9%; Score 223.2; DB 10; Length 616:  
 Best Local Similarity 93.5%; Pred No. 8.2e-44;  
 Matches 286; Conservative 0; Mismatches 15; Indels 5; Gaps 5;  
 Db 426 CCCGGGGATGCGCTGCTGGTACTCTGGCTGGGGATGGACATGATGAGGCCCTCC 436  
 Qy 437 CACCCCATTCCTCTGTGAGCAGGTGAAGSTCCCATAATGGAAACACATTGTGAGC 496  
 Db 486 CACCCCATTCCTCTGTGAGCAGGTGAAGSTCCCATAATGGAAACACATTGTGAGC 545  
 Qy 497 CAAATACACCTTGCGCTACAGGGAGACGAGTCGCATGTCCGTGAGACATGC 556  
 Db 546 CAATATACACCTTGCGCTACAGGGAGACGAGTCGCATGTCCGTGAGACATGC 605  
 Qy 557 TGTGCGCCGGAACACCGGAGGACTATGCCAGGGACTCGGAGGGCTTGGT 616  
 Db 606 TGTGCGCCGGAACACCGGAGGACTATGCCAGGGACTCGGAGGGCTTGGT 665  
 Qy 617 GCAAGGTGATGGCACTGCTGGAGGAGACGAGTCGCATGTCCGTGAGACATGC 676  
 Db 666 GCAGCGTGAATGGCACTGCTGGAGGAGGCTCGTCACTGGCGAGGGCTTGGCC 725  
 Qy 677 AGCCAACCGGCCCTGCACATCACCCGSGTCACTACTTGACTTGACTGATCCACCT 736  
 Db 726 AGCCAACCGGCCCTGCACATCACCCGSGTCACTACTTGACTTGACTGATCCACCT 785  
 Qy 737 ATGCCCCAAAAGCGTGAAGGGCC 763  
 Db 786 ATGCCCCAAAAGCGTGAAGC 812  
 RESULT 2  
 US-09-954-456-2242/C  
 ; Sequence 2242, Application US/09954456  
 ; Patent No. US20020115057A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Young, Paul  
 ; TITLE OF INVENTION: Process for Identifying Anti-cancer Therapeutic Agents Using Canc  
 ; TITLE OF INVENTION: Sets  
 ; FILE REFERENCE: 689290-76  
 ; CURRENT APPLICATION NUMBER: US/09/954, 456  
 ; CURRENT FILING DATE: 2001-09-18  
 ; PRIOR APPLICATION NUMBER: US/60/233, 617  
 ; PRIOR FILING DATE: 2000-09-18  
 ; PRIOR APPLICATION NUMBER: US/60/234, 052  
 ; PRIOR FILING DATE: 2000-09-20  
 ; PRIOR FILING DATE: US/60/234, 923  
 ; PRIOR FILING DATE: 2000-09-25  
 ; PRIOR APPLICATION NUMBER: US/60/235, 134  
 ; PRIOR FILING DATE: 2000-09-25  
 ;  
 ; US-09-594-456-2242  
 ; Sequence 1, Application US/10041006A  
 ; Patent No. US20020116875A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Andrade-Gordon, Patricia  
 ; APPLICANT: Darrode, Andrew  
 ; APPLICANT: Oi, Jian-shen  
 ; TITLE OF INVENTION: DNA encoding the novel human serine  
 ; TITLE OF INVENTION: Protease T  
 ; FILE REFERENCE: ORT-1032  
 ; CURRENT APPLICATION NUMBER: US/10/041, 006A  
 ; CURRENT FILING DATE: 2002-01-07  
 ; NUMBER OF SEQ ID NOS: 11  
 ; SOFTWARE: Patentin Ver. 2.0  
 ; SEQ ID NO: 1  
 ; LENGTH: 1110  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-10-041-006A-1

Query Match 26.1%; Score 201.4; DB 9; Length 1110;  
 Best Local Similarity 57.4%; Pred. No. 1.2e-38; Indels 301; Mismatches 0; Gaps 3; Matches 433; Conservative 0; Organism: Homo sapiens; SEQ ID NO 1

QY 12 GAAAGAATGTCGGGGTCAAGGAGCCCCAGGAGCAAGTGCCCTGGCAGGTAGCCT 71  
 Db 131 GAACCGAATGTCGGGGCAGGACACGGAGGAGGGAGGGAGTGGCCCTGGCAAAGTCAGCAT 190  
 QY 72 GAGACTCCAGGCCCATACTGATGACTGACTTCTGGGGGCTCCCTCATCACCCCACTG 131  
 Db 191 - - - - - CCAGGCCAACGGAAAGGCACTCTGGGGGAGGCCCATCGGGAGCAGCTG 241  
 QY 186 CAGGGTCAACTGGGAGGAGCACCTACTACAGGAGGACGCTGCTGGGGTAGCAGCAG 245  
 Db 302 GCTGGGGCAAGGCGCTAGTGACGGACACACGGTATGATGCCGGTGAGGA 361  
 QY 246 GATCAGTCGTCACCCACAGTCTACACCSCCAGATCGGAGGCAATGCCCTCTGTGA 305  
 Db 362 GGTGGAGGAAACCCCTGTACAGGGAGGGCTCCAGGGTAGGCTGAGTGGCCCTGGTGA 421  
 QY 306 GCTGGAGGAGCCGGCTGGGAGGTCTCCAGGACAGTCCAGGGTACCCCTGGCCTGCCC 365  
 Db 422 GCTGGAGGACCCAGTGCCTTCACCAATTACATCCTCCGGTGTGCTGACCCCTG 481  
 QY 366 AGAGACCTCCCCGGGGATGGCTGACTGGCTGGGATGTGGACATGA 425  
 Db 482 GGTGATCTTGTAGAGGGCATGAACTGCGUGGGTACTGTGGGAGGCCAGTGAGGA 541  
 QY 426 TGAGGCCCTCCACGCCATTCTCTGTGAAGCAGGCTGGTCACTGGAAACCA 485  
 Db 542 AGACCTCCGCCGACCCCTGTGACGGAGGCCCTGGGCTCAGGGTAGCTGGCTG 601  
 QY 486 CATTTGTGAGGCAAATAACACCTGGGCTACAGGGAGGACAGTCGGCATGGCTCC 545  
 Db 602 CAAGTGCACACTGCTCTACAGCAAGAACGACAGCTTGGCTTACCAACCATCA 661  
 QY 546 TGACGACAGCTGTGTCGGCG----GAACACGGGAGGACTATGCCAGGGGACTC 599  
 Db 662 GAATGACATGCTGTCGGCGCTTGAGGGAGGAAAGAGATGCTGAAAGGGGACTC 721  
 QY 600 CGGAGGGCCCTGGGTGCAAGGGAATGGCTGGCCTGGCTGGTCACTGGTCACTGGT 659  
 Db 722 GGCGGCCCTGGGTGCTGCGAGACGCCAGGTGTCTACATCCGCTGACGCCACCA 841  
 QY 720 GGACTGGATCACCACATGCCCCAAAAGCCGT 754  
 Db 842 CAACNGATCCATCGGATCATCCCCAACTGCAST 876

RESULT 4  
 US-10-040-655-1  
 ; Sequence 1, Application US/10040655  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Andrade-Gordon, Patricia  
 ; APPLICANT: Darrow, Andrew  
 ; TITLE OF INVENTION: DNA encoding the novel human serine  
 ; TITLE OF INVENTION: Protease T  
 ; FILE REFERENCE: ORT-1032  
 ; CURRENT APPLICATION NUMBER: US/10/040,655  
 ; CURRENT FILING DATE: 2002-01-07  
 ; NUMBER OF SEQ ID NOS: 11  
 ; SOFTWARE: PatentIn Ver. 2.0

RESULT 5  
 US-10-028-072-221  
 ; Sequence 221, Application US/10028072  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Beresini, Maureen  
 ; APPLICANT: Deforge, Laura  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Gao, Wei-Qiang

APPLICANT: Gerritsen, Mary E.  
 APPLICANT: Goddard, Audrey  
 APPLICANT: Godowski, Paul J.  
 APPLICANT: Gurney, Austin L.  
 APPLICANT: Sherwood, Steven  
 APPLICANT: Smith, Victoria  
 APPLICANT: Stewart, Timothy A.  
 APPLICANT: Tumas, Daniel  
 APPLICANT: Watanabe, Colin K  
 APPLICANT: Wood, William  
 APPLICANT: Zhang

TITLE OF INVENTION:  
 FILE REFERENCE:  
 CURRENT APPLICATION NUMBER: US/10/028,072  
 CURRENT FILING DATE: 2001-12-19  
 PRIOR APPLICATION NUMBER: 60/049911  
 PRIOR FILING DATE: 1997-06-18  
 PRIOR APPLICATION NUMBER: 60/056974  
 PRIOR FILING DATE: 1997-08-26  
 PRIOR APPLICATION NUMBER: 60/059113  
 PRIOR FILING DATE: 1997-09-17  
 PRIOR APPLICATION NUMBER: 60/059115  
 PRIOR FILING DATE: 1997-09-17  
 PRIOR APPLICATION NUMBER: 60/059117  
 PRIOR FILING DATE: 1997-09-17  
 PRIOR APPLICATION NUMBER: 60/059122  
 PRIOR FILING DATE: 1997-09-17  
 PRIOR APPLICATION NUMBER: 60/059184  
 PRIOR FILING DATE: 1997-09-17  
 PRIOR APPLICATION NUMBER: 60/059263  
 PRIOR FILING DATE: 1997-09-18  
 PRIOR APPLICATION NUMBER: 60/059352  
 PRIOR FILING DATE: 1997-09-19  
 PRIOR APPLICATION NUMBER: 60/059588  
 PRIOR FILING DATE: 1997-09-19  
 PRIOR APPLICATION NUMBER: 60/059836  
 PRIOR FILING DATE: 1997-10-17  
 PRIOR APPLICATION NUMBER: 60/062250  
 PRIOR FILING DATE: 1997-10-17  
 PRIOR APPLICATION NUMBER: 60/062285  
 PRIOR FILING DATE: 1997-10-17  
 PRIOR APPLICATION NUMBER: 60/062287  
 PRIOR FILING DATE: 1997-10-17  
 PRIOR APPLICATION NUMBER: 60/062814  
 PRIOR FILING DATE: 1997-10-24  
 PRIOR APPLICATION NUMBER: 60/062816  
 PRIOR FILING DATE: 1997-10-24  
 PRIOR APPLICATION NUMBER: 60/063045  
 PRIOR FILING DATE: 1997-10-24  
 PRIOR APPLICATION NUMBER: 60/063082  
 PRIOR FILING DATE: 1997-10-31  
 PRIOR APPLICATION NUMBER: 60/063127  
 PRIOR FILING DATE: 1997-10-24  
 PRIOR APPLICATION NUMBER: 60/063327  
 PRIOR FILING DATE: 1997-10-27  
 PRIOR APPLICATION NUMBER: 60/063329  
 PRIOR FILING DATE: 1997-10-27  
 PRIOR APPLICATION NUMBER: 60/063550  
 PRIOR FILING DATE: 1997-10-28  
 PRIOR APPLICATION NUMBER: 60/063561  
 PRIOR FILING DATE: 1997-10-28  
 PRIOR APPLICATION NUMBER: 60/063704  
 PRIOR FILING DATE: 1997-10-29  
 PRIOR APPLICATION NUMBER: 60/063733  
 PRIOR FILING DATE: 1997-10-29  
 PRIOR APPLICATION NUMBER: 60/063735  
 PRIOR FILING DATE: 1997-10-29  
 PRIOR APPLICATION NUMBER: 60/063738  
 PRIOR FILING DATE: 1997-10-29  
 PRIOR APPLICATION NUMBER: 60/063755  
 PRIOR FILING DATE: 1997-10-17  
 PRIOR APPLICATION NUMBER: 60/064248  
 PRIOR FILING DATE: 1997-11-03  
 PRIOR APPLICATION NUMBER: 60/066770  
 PRIOR FILING DATE: 1997-11-24  
 PRIOR APPLICATION NUMBER: 60/069334  
 PRIOR FILING DATE: 1997-12-11  
 PRIOR APPLICATION NUMBER: 60/069964  
 PRIOR FILING DATE: 1997-12-16  
 PRIOR APPLICATION NUMBER: 60/072320  
 PRIOR FILING DATE: 1998-01-23  
 PRIOR APPLICATION NUMBER: 60/077991  
 PRIOR FILING DATE: 1998-02-04  
 PRIOR APPLICATION NUMBER: 60/078910  
 PRIOR FILING DATE: 1998-02-09  
 PRIOR APPLICATION NUMBER: 60/079409  
 PRIOR FILING DATE: 1998-03-25  
 PRIOR APPLICATION NUMBER: 60/079663  
 PRIOR FILING DATE: 1998-02-27  
 PRIOR APPLICATION NUMBER: 60/079728  
 PRIOR FILING DATE: 1998-03-27  
 PRIOR APPLICATION NUMBER: 60/079294  
 PRIOR FILING DATE: 1998-03-31  
 PRIOR APPLICATION NUMBER: 60/081203  
 PRIOR FILING DATE: 1998-04-09  
 PRIOR APPLICATION NUMBER: 60/081229  
 PRIOR FILING DATE: 1998-04-09  
 PRIOR APPLICATION NUMBER: 60/081695  
 PRIOR FILING DATE: 1998-04-14  
 PRIOR APPLICATION NUMBER: 60/081817  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/081818  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/082999  
 PRIOR FILING DATE: 1998-04-24  
 PRIOR APPLICATION NUMBER: 60/083322  
 PRIOR FILING DATE: 1998-04-28  
 PRIOR APPLICATION NUMBER: 60/083545  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/084600  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084627  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084637  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/085149  
 PRIOR FILING DATE: 1998-05-12  
 PRIOR APPLICATION NUMBER: 60/085323  
 PRIOR FILING DATE: 1998-05-13  
 PRIOR APPLICATION NUMBER: 60/085338  
 PRIOR FILING DATE: 1998-05-13  
 PRIOR APPLICATION NUMBER: 60/085339  
 PRIOR FILING DATE: 1998-05-13  
 PRIOR APPLICATION NUMBER: 60/085579

PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085697  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085704  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/086414  
PRIOR FILING DATE: 1998-05-22  
PRIOR APPLICATION NUMBER: 60/086430  
PRIOR FILING DATE: 1998-05-22  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-05-04  
PRIOR APPLICATION NUMBER: 60/088730  
PRIOR FILING DATE: 1998-05-10  
PRIOR APPLICATION NUMBER: 60/088741  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/099947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090445  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090538  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090863  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/091360  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091519  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091982  
PRIOR FILING DATE: 1998-07-07

Query Match 26.1%; Score 201.4; DB 9; Length 1129;  
Best Local Similarity 57.4%; Pred. No. 1.2e-38;  
Matches 433; Conservative 0; Mismatches 301; Indels 21; Gaps 3;

Qy 12 GAAAGAATGTCGGGGTCAAGGAGGCCAGGGCAAGTGCCCTGGCAGGTGCCT 71  
Db 131 GAACCGAATGCTGGGGCAGGACACGGAGGGCAGTGCCCTGGCAAGTCAGCAT 190

Qy 72 GAGAGTCAGGCCATACTGATGCACTCTGGGGGCTCCCATCACCCAGTG 131  
Db 191 -----CCAGCCACAACGAAGCACTCTGCTGGGGCAGCCATCGGGAGCATG 241

Qy 132 GGTGGTGACGCCGGGCTGGCAACTCTGGCTGGCAACGATCTGGCC----GCCCT 185  
Db 242 GGTGGTGACGCCGGGCTGGCAACTCTGGCTGGCAACGATCTGGCC----GCCCT 301

Qy 186 CAGGGGCAACTGGGAGGAGCACCTCTACTACAGGACCGCTGCTGGGGTAGCAG 245  
Db 302 GCTGGGGCAAGGGCACTGTGAGGGGACACACGCATATGTTAGCCGGTAGGCA 361

Qy 246 GATCATGTCGACCCAGCTTACACGCCAGATGAGGAGGAGCATGGCCCTGG 305  
Db 362 GGTGGAGGAGAACCCCTGTACAGGGCAAGGCCCTCAGGGCTGAGGGCTGG 421

Qy 306 GCTGGAGGCCGGGAGAGTCTCCAGCACGGTACCCGGCCCTGGCT 365  
Db 422 GCTGGAGGCCGGCAGCTAGTGAGGGGACACAGGCTATGTATGCCCGGTGAGG 481

Qy 366 AGAGACCTCCCCGGGATGCCGTGCTGGCACTGCTGCTGGGGCAGTGGACAATGA 425  
Db 482 GGTGATCTTGTAGACGGGATGAACTGCTGGTCACTGCTGGGGCAGCCCAGTGAGGA 541  
Qy 426 TGAGGCCCTCCACGCCATTCTCTGAGACAGTGGAGGTGCCATATGAAANCCA 485  
Db 542 AGACCTCCCTGCCAGCCGGATCTGAGAACTCGCTGTGCCATCATCGACACCC 601  
Qy 486 CATTTGTGACCAAATACACCTGGGCCATACCGGGAGACGACGTCGGCATGTCG 545  
Db 602 CAAGTGCAAGCTGCTACAGCAAGAACAGCTGGGACTTGTGCTCACCCAAACCATCAA 661  
Qy 546 TGACGACATGCTGTGTCGGG-----GACACCGGAGGACTCATGCCAGGGGACTC 599  
Db 662 GATGACATCTGTCGGCCGGCTGAGGGGGAAAGAGGATGCCATCGACGACTC 721  
Qy 600 CGGAGGCCCTGGTGCAGAGGAATGCGACCTGGCTGCAAGGGGGTGTGTCASGTG 659  
Db 722 GGGGGCCCTGGTGCCTCGTGGCTAGTCGTCGGCAGGGGGGTGATCAGCTG 781  
Qy 660 GGGGAGGGCTGTGCCAGGCCAACCGGGCTGGCATCTACCCGGTGTACACTACTT 719  
Db 782 GGTTGAGGGGTGTCGGCGCAGAACGGCCAGGTTGTCATCGGTGTCACCCACCA 841  
Qy 720 GAGCTGGATCCACCATATGCCAAAAAGCCGT 754  
Db 842 CAACTGGATCCATCGGATCATCCCAACTGCGATG 876

RESULT 6  
US-10-041-006A-8  
Sequence 8, Application US/10041006A  
Patent No. US2002168754A1  
GENERAL INFORMATION:  
APPLICANT: Andrade-Gordon, Patricia  
APPLICANT: Darrow, Andrew  
APPLICANT: QI, Jian-shen  
TITLE OF INVENTION: DNA encoding the novel human serine protease  
FILE REFERENCE: ORT-1032  
CURRENT APPLICATION NUMBER: US/10/041, 006A  
CURRENT FILING DATE: 2002-01-07  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 8  
LENGTH: 1130  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Fusion gene of  
OTHER INFORMATION: Protease T in a zymogen activation vector  
US-10-041-006A-8

Query Match 25.1%; Score 193.8; DB 9; Length 1130;  
Best Local Similarity 56.7%; Pred. No. 7.5e-37;  
Matches 429; Conservative 0; Mismatches 307; Indels 21; Gaps 3;

Qy 10 GAGAAAGAATGTCGGGGCTAGGAGGCCAGGAGGTGGCCCTGGCAGGTGAGC 69  
Db 157 GATGACAAGTCGTTGGGGCTATGCTCTAGAGGAGGGAGTGCCCCCTGGCAGTCAG 216  
Qy 70 CTGAGAGTCACGGCCATACTGGTCACTCTGGGGGCTCCCTCATCACCCCG 129  
Db 217 AT-----CCAGGCACAGGAAGCCACTCTCGGGGGCAGGCCATCGGGAGCAG 267

Qy 130 TGGGGCTGACGCCGGGGCTGGGAGGAGCTGGCTGGGGCTGG-----CCGC 183  
Db 268 TGGGCTCTGAGGGCTGGCACTCTGGCTCCGCAACACCTCTGAGGCTCCCTGTACCGAG 327

Qy 184 CTCAAGGTGCACTGGGGAGGAGCACCTACTACAGGACCTGTGCGGGTCACT 243  
Db 328 CTGCTGGGGCAAGGCACTGTGAGGGGACACAGGCTATGTATGCCCGGTGAGG 387

QY 244 AGCATCATCGTCACCCACAGTCTCACCCGCCAGATCGGAGCAGATCCGCCCTCTG 303  
 Db 388 CAGGTGGAGGACACCCCTGTACCAAGGCACGCCCTCAGGGTGAAGTGGGCTGGT 447

QY 304 GACCTGGAGGACCGCGTGAAGGCTCTCAGGCCACGGTCAACCGTCACTGCCCAGCC 363  
 Db 448 GAGCTGGAGGACCGTGTGCCCTACCCATTACATCCCTCCCCGTGTCAGCCC 507

QY 364 TCGAGACCTCCCCGGGATGCCGCTGCTGGTACTGGTCACTGGTGGGGATGTGGCAAT 423  
 Db 508 TCGTGAATCTTGAGAAGCAGTGACTGCTGGTACTGGTCACTGGGGCAGCCAG 567

QY 424 GATGAGCCTCCACCCGCTATGGAAAC 483  
 Db 568 GAGAGCTCTGCCAACCGGGATCTGAGAAACTCGCTGCTGCCATCATCGACACA 627

QY 484 CACATTGCTGACCACAAATACACCTTGCGCTACAGGGAGACAGGAGTCCAGG 543  
 Db 628 CCCAGTGCACCTGCTACAGCAAGACACCCGACTATGCCAGGGGAC 687

QY 544 CGTGACGACATGCTTGCGCG----GAACACCCGAGGACTATGCCAGGGGAC 597  
 Db 688 AAGATGACATGCTTGCGCGCTTCCAGGAAGACACCCAAACCAAC 747

QY 598 TCCGGAGGGCCCTGGTGTCAAGGTGATGGCACTGGCCTGCGAGGGCGTGGTCAGC 657  
 Db 748 TCGGGCGGCCCTGGTGTGGCTCTGCTGAGGGGGTGTAGC 807

QY 658 TGGGGCGAGGCTGTGCCAACCGGGCTGTGCACTTACACCCGTTGAC 717  
 Db 808 TGGGTGAGGGCTGTGCCAACCGGGCTGTACATCCGGTCACTGCCAAC 867

QY 718 TTGACTGATCCACCATATGCCCCAAAAGCGT 754  
 Db 868 CACACTGGATCCATCGGATCATCCCCAACTGGAGT 904

RESULT 7  
 US-10-040-655-8  
 ; Sequence 8, Application US/10040655  
 ; Patent No. US20020146805A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Andrade-Gordon, Patricia  
 ; APPLICANT: Darrow, Andrew  
 ; APPLICANT: Qiu, Jian-shen  
 ; TITLE OF INVENTION: DNA encoding the novel human serine  
 ; TITLE OF INVENTION: Protease T  
 ; FILE REFERENCE: ORT-1032  
 ; CURRENT APPLICATION NUMBER: US/10/040,655  
 ; CURRENT FILING DATE: 2002-01-07  
 ; NUMBER OF SEQ ID NOS: 11  
 ; SOFTWARE: Patentin Ver. 2.0  
 ; SEQ ID NO 8  
 ; LENGTH: 1130  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: Fusion gene of  
 ; OTHER INFORMATION: Protease T in a zymogen activation vector  
 ; US-10-040-655-8

Query Match 25.1%; Score 193.8; DB 12; Length 1130;  
 Best Local Similarity 56.7%; Pred. No. 7.5e-37; Mismatches 307; Matches 429; Conservative 0; Mismatches 307; Indels 21; Gaps 3;

QY 10 GAGAAAGAAGATCGTGGGGTCAGAGGCCAACAGCAAGTGGCCCTGGCAGTGAGC 69  
 Db 157 GATGACAGATCGTGGGGTATGCTCTAGAGGGGGAGTGGCCCTGGCAAGTCAGC 216

QY 70 CTGAGAAGTCCACGCCATACTGGATGCACTTCCTGGGGCTCCTCATCCACCCCAAG 129  
 Db 217 AT-----CCACCGCAACCGGAAGCCAC-----CTGGGGGCAGCTCATGCCAG 267

QY 130 TGGTGTGACCCCGGGCGCGGGAGGAGCTCAAGGAATCTGG----CCGCC 183  
 Db 268 TGGTCTGTGACGCGCTCGCACTGCTTCGGCACACCTCTGAGACGTCCTGTACAGTC 327

QY 184 CTAGGGTGCAACTGCGGGAGGAGCACCTCTACCTACAGGAGCAGCTGCTCCGGTCAGC 243  
 Db 328 CTCTGGGGCAGGAGCTAGTGCAGCGGGACACGGTATGATGCCGGTGAG 387

QY 244 AGATCATATGTCACCCACAGTCTACCCGCAAGATGGGCACTGGGCAAT 303  
 Db 388 CAGTGGAGGAAACCCCTGACAGGGCAGGGCTACGGCTACGGCTACGGCTACGGCCCTGGG 447

QY 304 GAGCTGGAGGACCCGCGTCAAGGGTGAAGGCTCCATAATGGAAAC 483  
 Db 448 GAGCTGGAGGACACAGTGCCTCTCACAAATTACATCCCTCCGGTGTCCCTGCTGACCCC 507

QY 364 TAGAGACCTCCCCGGGATGCCGCTGTGAAGCAGGACTATGGGGGAGATGGCAAT 423  
 Db 508 TCGGTGACCTTGAGAGGGCTATGCTGCTGGTCACTGGGGGAGATGGCAAC 567

QY 424 GATGAGCCTCTGCCAACCGGGCTTCCCTGTAAGGCACTGCTGGTGGGGCAGCCAG 483  
 Db 568 GAGAGCTCTGCCAACCGGGCTTCCCTGCAAGGAGATGCTGAGGGCAGCCAG 627

QY 484 CACATTGCTGACCACAAATACACCTTGCGCTACAGGGAGACAGGAGTCCAGG 543  
 Db 628 CCCAGTGCACCTGCTACAGCAAGACACCCGACTATGCCAGGGGAC 687

QY 544 CGTGACGACATGCTTGCGCG----GAACACCCGAGGACTATGCCAGGGGAC 597  
 Db 688 AAGATGACATGCTTGCGCGCTTCCAGGAAGACACCCAAAC 747

QY 658 TGGGGCGGCCCTGGTGTGGCAGAGGGGCTGTGAGGAGGATGTCAGGAGCACA 807

QY 718 TTGACTGATCCACCATATGCCCCAAAAGCGT 754  
 Db 868 CACACTGGATCCATCGGATCATCCCCAACTGGAGT 904

RESULT 8  
 US-09-888-615-35  
 ; Sequence 35, Application US/09888615  
 ; Patent No. US2002006856A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: PLOWMAN, GREGORY  
 ; APPLICANT: WHYTE, DAVID  
 ; APPLICANT: CAENEPEEL, SEAN  
 ; APPLICANT: CHARNDZAK, GLEN  
 ; APPLICANT: MANNING, GERARD  
 ; APPLICANT: SUDARSHAN, SUCHA  
 ; TITLE OF INVENTION: NOVEL PROTEASES  
 ; FILE REFERENCE: 038602/1214  
 ; CURRENT APPLICATION NUMBER: US/09/888, 615  
 ; CURRENT FILING DATE: 2001-05-26  
 ; PRIOR APPLICATION NUMBER: 60/214, 047  
 ; NUMBER OF SEQ ID NOS: 150  
 ; SOFTWARE: Patentin Ver. 2.1  
 ; SEQ ID NO 35  
 ; LENGTH: 2847  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens

Query Match 24.9%; Score 191.6; DB 10; Length 2847;  
 Best Local Similarity 59.5%; Pred. No. 3e-36;

Matches	370;	Conservative	0;	Mismatches	234;	Indels	18;	Gaps	2;
Qy	141	CGCCGCGCGCTGGCTGGCTGGACCGAACGTCAGGTCAAGGA	-	T	TGGCCGCCTCAGGTGAACTCGC	200			
Db	612	CCCCCGCCCTGCA	GGGAGGAGAGTGGAGGCTTGAGATGTCAGGCTGG	671					
Qy	201	GGAGCAGGACCTCTACTACCAGAACAGCTGCTGCCGTCAGGAG	CATGTCGACCC	260					
Db	672	GCACTGAGGCTATGAGGACGACAGCGAACAGGAGTGGAGATGTC	CGCTACCC	731					
Qy	261	ACAGTCTACA-----CCGAGCATGGAGGGACATGCC	CTGGAGCTGGA	311					
Db	732	CCAGTACACAGAGGCCUGTCCTCCAGGCGCTGCGACATGCC	CGCTGAGCTGA	791					
Qy	312	GGAGCGGTGAAGGTTCTCAGCCAGGTCACACGGTCACCC	TGCTGCCCTGAGAGAC	371					
Db	792	GGCCCGGTGCCGTCTGAGCTCATCCACCGGTCTCGCTCC	TGCTGCGCTCC	851					
Qy	372	CTTCCCCCGGGATGCCGTGCTGCCGACTSCTGSGGGA	GTGAGCTGGA	431					
Db	852	CGTCCCCTGGGAGACCTGCTGCTGACCGGGTGGTCAATTG	GAGA	911					
Qy	432	CCTCCACGCCATTCTCTGAAGCAGTGAAGGCCATATG	GAACACATTG	491					
Db	912	ACTGCCCTGCCCTCACCTGTRGGAGCAGGTGAGGTCA	GGACCAAGCTCTG	971					
Qy	492	TGACGCCAAATAACACAC-----CTTGGCGCTACAGG	GGAGACGAGCTCGCATGT	542					
Db	972	TAACCGAACCTGTCGCCGCCCTTCTCCAACCAACACTG	AGGGTTGAGGGCTCAT	1031					
Qy	543	CCCTGACCACTGCTGCTGCCGAAACCCGAGACTCATG	CCAGGGACTCG	602					
Db	1032	CAAGGACGACATGCTGTCGCGGGACCGACCGCTCG	CCAGGACACAG	1091					
Qy	603	AGGCCCTGGTGTCAAGGTGAATGGACTCGCTG	AGGGGGGTGTCAGCTGG	662					
Db	1092	GGCCCCCTCCTGAGCGGGAAATGCACTGGTCAAGG	GGGGACTCG	1151					
Qy	663	CGAGGCCTGGTGCACGCCAACCGGCCCTGGCATCTAC	ACCCGGTGTACCTACTACTTG	722					
Db	1152	CAARACTCTGGCTTCGGCTATCCGGCATGTACACCC	GGTGTACGGACTACG	1211					
Qy	723	CTGGATCCACCACTATGRC	CCC 744						
Db	1212	CTGGATCCGCCAGCCATGCC	1233						
<b>RESULT 9</b>									
us-10-041-400A-1									
; Sequence 1: Application US/10041400A									
; Patent No. US2002010895A1									
; GENERAL INFORMATION:									
; APPLICANT: Darrown, Andrew									
; APPLICANT: Andrade-Gordon, Patricia									
; APPLICANT: QJ, Jenson									
; TITLE OF INVENTION: DNA Encoding the Human Serine									
; TITLE OF INVENTION: Protease BOS									
; FILE REFERENCE: ORT-1031									
; CURRENT APPLICATION NUMBER: US/10/041,400A									
; CURRENT FILING DATE: 2002-01-08									
; PRIOR APPLICATION NUMBER: US/09/387,375									
; PRIOR FILING DATE: 1999-08-31									
; NUMBER OF SEQ ID NOS: 9									
; SOFTWARE: Patentin Ver. 2.0									
; SEQ ID NO 1									
; LENGTH: 1613									
; TYPE: DNA									
; ORGANISM: Homo sapiens									
; US-10-041-400A-1									
Query Match 24.4%; Score 187.8%; DB 12; Length 1613;									
Best Local Similarity 55.9%; Pred. No. 2.1e-35; Mismatches 292; Indels 21; Gaps 3;									
Qy	17	GAATCGTGGGGTCAGGAGCCCCAGGAGCAAGTGGCCC	TGAGAG	76					
Db	175	GATCTGTTGGGGCGGGATGCCGAGGAGTGGCGCTG	CGAGCATCCAGC	234					
Qy	77	TCCACGGCCCATACTGGATGCACTCTGGGGG	CTCTCAGCCCAC	136					
Db	235	ATCTCTGGGCA-----CACGTGTCGGGG	GGCTCATGCC	285					
Qy	137	TGACGGCCGGCTGCTGGACCGGGACGTCAAG-----	GATCTGGCCCTCAGG	190					
Db	286	TGACAGCSCGCACTGCTTCCCAGGGAGCTGCA	GCTTGGCGCTTGG	345					
Qy	191	TGCAACTGGGGACGACCTCTACTACAGGACACGCTG	CGCTGCGCTCAGGAGATCA	250					
Db	345	GGGGCTGCTGCTGGCTCAGGCTCCCTGCGCC	CGACGGTTC	405					
Qy	251	TGCTGACACAGTTCTCACACGCCAGATCGAGCGAC	ATGCGCTGCTGAGCTG	310					
Db	406	TGCTGCCCCGGACTACTCCGAGGACGGGGAC	TGCTGAGCTG	465					
Qy	311	AGAGCCGGTAGGTCAGCTGCCACAGCTGAGG	CCCTGCTCCTGCTCAGAGA	370					
Db	466	GTCCGCCGCGCCGCTGCGCCGAGTCGAGG	GGCGCTGCGCC	525					
Qy	371	CTTCCCCCGGGATGCCGTGGCTGCGGTCAC	GCTGAGCTGAG	430					
Db	525	GCCGGCGCCGGCACACCATGCCGGTCA	CCCTGGGCGACCTCCG	585					
Qy	431	GCCTCCACGCCATTCTCTGAGCTGAGGCTTATG	GAACACATTG	490					
Db	586	CCCTCCAGAGTGGCGACGCCGCTACAGGAGT	AGGGTGGCTG	645					
Qy	491	GTGAGCCTAACACACTTGGGCCACACGGAGAC	GACCTCCGGCTG	550					
Db	646	GCAGGGCTCTTACACCTGGGGGGAGCTCC	GGCGCC	705					
Qy	551	ACATGCTGTTGGGGACACCCG-----AGGGACT	CATGCCAGGGGACTCCGG	604					
Db	705	GGACTCTGCTGTCGGCTCACCC	GGGGCACAGGAGCCTCCAGGTGATCTGG	765					
Qy	605	GGCCCTGGTGCAGGTTGACATGGGACTCG	GGCGCGCGTGTG	664					
Db	766	GACCTCTGACCTGCTGGAGCTGGGTCTGG	GGGTGAGCTGGGGCA	825					
Qy	665	AGGCTGTCGCCAGCCCACGGCCTGCATCTACAC	CGTGTGAGCGCGCGCGTGTGACTTGTGACT	724					
Db	826	GGTTGCGCCCTGCCACGCTTACACGGTGTG	GGCCACATGTAGCCCT	885					
Qy	725	GGATCCA	731						
Db	886	GGATTCA	892						
<b>RESULT 10</b>									
US-10-041-264A-1									
; Sequence 1: Application US/10041264A									
; Patent No. US2002012446A1									
; GENERAL INFORMATION:									
; APPLICANT: Darrown, Andrew									
; APPLICANT: Andrade-Gordon, Patricia									
; APPLICANT: QJ, Jenson									
; TITLE OF INVENTION: DNA Encoding the Human Serine									
; TITLE OF INVENTION: Protease BOS									
; FILE REFERENCE: ORT-1031									
; CURRENT APPLICATION NUMBER: US/10/041,64A									
; CURRENT FILING DATE: 2002-01-08									
; PRIOR APPLICATION NUMBER: US/09/387,375									
; PRIOR FILING DATE: 1999-08-31									
; NUMBER OF SEQ ID NOS: 9									
; SOFTWARE: Patentin Ver. 2.0									
; SEQ ID NO 1									
; LENGTH: 1613									
; TYPE: DNA									
; ORGANISM: Homo sapiens									
; US-10-041-264A-1									



APPLICANT: Allen, Keith D.  
 APPLICANT: Levitin, Michael W.  
 TITLE OF INVENTION: TRANSGENIC NICE CONTAINING TRYPTHASE GENE  
 FILE REFERENCE: R-372  
 CURRENT APPLICATION NUMBER: US/09/900,754  
 CURRENT FILING DATE: 2001-07-06  
 PRIOR APPLICATION NUMBER: US 60/216,109  
 PRIOR FILING DATE: 2000-07-06  
 PRIORITY NUMBER: US 60/223,172  
 PRIORITY NUMBER: US 60/244,111  
 PRIORITY FILING DATE: 2000-08-07  
 SEQ ID NO: 1  
 SOFTWARE: FastSEQ for Windows Version 4.0  
 LENGTH: 1122  
 TYPE: DNA  
 ; ORGANISM: Mus musculus  
 ; US-09-900-754-1

Query Match 24.3%; Score 187.6; DB 10; Length 1122;  
 Best Local Similarity 58.9%; Pred. No. 2.1e-35; Mismatches 0;  
 Matches 432; Conservative 0; Mismatches 269; Indels 33; Gaps 5;

QY 17 GAATGCTCGGGGTAGGAGCCCCAGGAGCAAGTGGCCCTGGAGGTGAGCTGAGAG 76  
 Db 86 GAATGCTGGAGGGCATGCCAACAGCAGGCACATTGGCTGGAGGTAGCT---- 140  
 Qy 77 TCCACGGCCATACTGGTCACTTCTGCGGGGGTCCCTCATCCACCCCACTGGTGC 136  
 Db 141 ---CGCTCAGGAGCTTGCGACACTGCTGCGGGCTCCCTGCTCAGTCAGGGCAAC 196  
 Qy 137 TGACCGCCGGGGCTGGGGCTGGGGACCTCAGATCGGCCCTCAGGGCAAC 196  
 Db 197 TCACAGCAGCCACTCTCTGCTGCTGCTGAGTC---GTCCTGTTATCAGGGCACT 253  
 Qy 197 TCGGGGAGCAGCACCTCTACCCAGGACCTCTGCGGTCCAGGATCATCGTC 256  
 Db 254 TGGGAGAGCTTACGGTCACACTGTCGCCACTTCTCCACTGTAAACGGATCATGT 313  
 Qy 257 ACCCAGAGTCTAC---ACGCCCAAGATCGAGGAGACATGCCCTGCTGAGCTGGAGG 313  
 Db 314 ACACGGCTCTCCAGGACACGGGGTCCAGTGGACATGCCCTGCTGAGCTGGCT 373  
 Qy 314 AGCCGGTAAGGTCAGCAGGTCACGGTACCCGCCCTGCTGAGCTAGACCT 373  
 Db 374 CCCCGTGGCCCTTCCAGCAGGCTGCTGCTGCTGCCAGGGCTCAGGTGACT 433  
 Qy 374 TCCCCCGGGATGCGCTGCGGACTCGCTGGCGATGTGGACAATGATGGGCC 433  
 Db 434 TCTACCTCTGGATGAGTGTGGTACTGCTGCTGGCTATACAGGGAGGAGCCTC 493  
 Qy 434 TCCCACGCCATTCTCTGAAGAATGCTCCATAATGGAAACACATTTGTG 493  
 Db 494 TGAACCCCCATACACCTTCAGGAGGCCAACGCTCTGTTGAT----- 540  
 Qy 494 ACGCAAATACACCTTGGCTCACGGAGAGACGTCGCCCTGCTGAGGACA 553  
 Db 541 -GTAAGACCTGCAAGCAGCTTACAAATGTCATGCCATGGCAGCTCATCGCAAGACA 598  
 Qy 554 TGCTGTTGCGGGACACCGGGAGGGACTCATGCCAGGCCACTCCGGGGCTGG 613  
 Db 599 TGCTATGCGCCGGGCCCTG---GGATGCTGCTGCCAGGATGACTCTGGAGGGCCACTAG 655  
 Qy 614 TGTGAGGAGATGGACCTGGCTGCGAGGGGGTGGTAGCTGGGGAGGGCTGTG 673  
 Db 656 TCTGCAACGGCTGGCATCTACCCGGTGTACCTACTGGACTGAGTGGATCCACC 715  
 Qy 674 CCCACCCAACGGGCTGGCATCTACCCGGTGTACCTACTGGACTGAGTGGATCCACC 733  
 Db 716 GCGGCGCTGACGCCCTGGGTCTGGCTATGCGGGGTACTGCTATGAACTGGATCCACC 775  
 ; ORGANISM: Artificial Sequence

RESULT 13  
 US-09-960-352-1410/c  
 Sequence 1410, Application US/09960352  
 Patent No. US20020137139A1  
 GENERAL INFORMATION:  
 APPLICANT: Warren, Wesley C.  
 APPLICANT: Tao, Mengping  
 APPLICANT: Byatt, John C.  
 APPLICANT: Mathialagan, Nagappan  
 TITLE OF INVENTION: NUCLEAR ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND MUSCLE AND FAT DEPOSITION  
 FILE REFERENCE: 16511.00673721(10290)C  
 CURRENT APPLICATION NUMBER: US/09/960,352  
 CURRENT FILING DATE: 2001-09-24  
 NUMBER OF SEQ ID NOS: 15112  
 SEQ ID NO: 1410  
 LENGTH: 249  
 TYPE: DNA  
 ; ORGANISM: Bos taurus  
 ; OTHER INFORMATION: Clone ID: 07-LIB188-011-Q1-E1-B11  
 ; US-09-960-352-1410  
 Query Match 23.9%; Score 184.4; DB 10; Length 249;  
 Best Local Similarity 85.1%; Pred. No. 8.7e-35; Mismatches 0;  
 Matches 206; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 503 ACCACCTGGGCCATACAGGGAGACGCTCGGATCGTCGCTGAGCATGCTGTG 562  
 Db 249 ACCACTCTGGCTGTCACCGGGGACAACCTCCCATCGGGGAGGACATGCTGTG 190  
 Qy 563 CGGGACACCCGGAGGACTCATGCCAGGGCACTCCGGGCCCCTGGTGTGCAAG 622  
 Db 189 CTGGGAGCAGGGAGGAACCTCTGCGCAGGGCAGCTCTGGGGCCCTGTGCAAGG 130  
 Qy 623 TGAATGGCACCTGGCTGCAAGGGGGCTGTCAGCTGGGGAGGGCTGGCCAGGCCA 682  
 Db 129 TGAATGGCACCTGGCTGCAAGGGGGCTGTCAGCTGGGGAGGGCTGGGAGGCCA 70  
 Qy 683 ACCGGCTGGCATCTACACCGGTGACCTACTGACTCTGGACTGACCCACTATGTC 742  
 Db 69 ACCGGCCGGCATCTACACCGGGTCACTCTACTGGACTGGATTCAGCGTATCCCG 10  
 Qy 743 CC 744  
 Db 9 CC 8

RESULT 14  
 US-10-041-400A-8  
 Sequence 8, Application US/10041400A  
 Patent No. US2002010895A1  
 GENERAL INFORMATION:  
 APPLICANT: Barrow, Andrew  
 APPLICANT: Andrade-Gordon, Patricia  
 APPLICANT: Oli, Jensen  
 TITLE OF INVENTION: DNA Encoding the Human Serine Protease EOS  
 FILE REFERENCE: ORT-1031  
 CURRENT APPLICATION NUMBER: US/10/041,400A  
 CURRENT FILING DATE: 2002-01-08  
 PRIORITY NUMBER: US/09/387,375  
 PRIORITY NUMBER: US/09/387,375  
 PRIORITY FILING DATE: 1999-08-31  
 NUMBER OF SEQ ID NOS: 9  
 SOFTWARE: Patentin Ver. 2.0  
 SEQ ID NO: 8  
 LENGTH: 1130  
 TYPE: DNA



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: January 31, 2003, 03:51:33 ; Search time 56 Seconds

4222.282 Million cell updates/sec

Title: US-09-598-982-20  
Perfect score: 771  
Sequence: 1 gggccctcgagaaaaaat.....cgtgaaggccgcgtcgt 771  
Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Post-processing: Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Listing first 45 summaries

Database : Issued Patents\_NA,\*

1: /cgn2\_6/ptodata/1/ina/5A-COMB.seq: \*  
2: /cgn2\_6/ptodata/1/ina/5B-COMB.seq: \*  
3: /cgn2\_6/ptodata/1/ina/6A-COMB.seq: \*  
4: /cgn2\_6/ptodata/1/ina/6B-COMB.seq: \*  
5: /cgn2\_6/ptodata/1/ina/6C-COMB.seq: \*  
6: /cgn2\_6/ptodata/1/ina/hackfille1.seq: \*

## SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## ALIGNMENTS

Result No.	Score	Query Match Length	DB ID	Description
1	764.6	99.2	771 4	US-09-079-970A-4
2	731	94.8	1128 2	US-09-016-366A-20
3	731	94.8	1128 2	US-09-078-404B-15
4	731	94.8	1137 2	US-09-016-366A-18
5	731	94.8	1137 2	US-08-978-404B-13
6	728.6	94.5	735 4	US-09-079-970A-1
7	726.2	94.2	1081 2	US-09-016-366A-22
8	726.2	94.2	1081 2	US-08-978-404B-17
9	686.2	89.0	1154 2	US-09-016-366A-16
10	686.2	89.0	1154 2	US-08-978-404B-11
11	477.2	61.9	1219 2	US-08-978-404B-7
12	469.6	60.9	1108 2	US-09-016-366A-14
13	469.6	60.9	1108 2	US-08-978-404B-20
14	463	60.1	1031 2	US-08-978-404B-1
15	445.6	57.8	1103 2	US-09-016-366A-24
16	444	57.6	1097 2	US-08-978-404B-4
17	371	48.1	2259 2	US-08-845-998-3
18	371	48.1	2259 3	US-09-206-537-3
19	371	48.1	2259 4	US-09-430-854-3
20	369.4	47.9	2218 2	US-08-845-998-5
21	369.4	47.9	2218 3	US-09-206-537-5
22	369.4	47.9	2218 4	US-09-430-854-5
23	271.8	35.3	1095 2	US-08-978-404B-9
24	261.4	26.1	1110 4	US-09-386-653A-1
25	193.8	25.1	1130 4	US-09-386-653A-8
26	153	19.8	3757 2	US-09-016-366A-13
27	3757	2	US-08-978-404B-19	
28	143.8	18.7	2397 2	US-08-978-404B-2
29	140.6	18.2	1142 4	US-09-386-642-8
30	140.6	18.2	1169 4	US-09-386-642-7
31	131.2	17.0	1430 4	US-09-386-629-1
32	129.2	16.8	1166 4	US-09-386-629-2
33	115.4	15.0	1081 4	US-09-008-271A-15
34	115.4	15.0	1103 4	US-09-386-642-59
35	112.8	14.6	2413 4	US-09-518-046-1
36	108.5	14.1	1109 4	US-09-088-651-6
37	107.6	14.0	1109 4	US-09-088-651-1
38	103.6	13.4	1386 2	US-08-756-506-3
39	103.5	13.4	1755 6	5225537-1
40	102.8	13.3	2416 4	US-09-261-416-1
41	100.8	13.1	654 1	US-09-280-116-57
42	98.4	12.8	2544 4	US-09-518-046-3
43	95.6	12.4	1155 4	US-09-163-951-15
44	95.6	12.4	1240 4	US-09-163-951-14
45	94.2	12.2	1387 6	5270178-1

Sequence 2, Appli  
Sequence 8, Appli  
Sequence 7, Appli  
Sequence 1, Appli  
Sequence 2, Appli  
Sequence 3, Appli  
Sequence 15, Appli  
Sequence 59, Appli  
Sequence 1, Appli  
Sequence 6, Appli  
Sequence 1, Appli  
Sequence 3, Appli  
Sequence 14, Appli  
Patent No. 525537  
Sequence 1, Appli  
Sequence 57, Appli  
Sequence 3, Appli  
Sequence 15, Appli  
Sequence 14, Appli  
Patent No. 5270178



**RESULT 3**

US-08-978-404B-15

Sequence 15, Application US/08978404B

Patent No. 598782

**GENERAL INFORMATION:**

APPLICANT: Stevens, Richard L.

TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES TITLE OF INVENTION: FIBRINOGEN

NUMBER OF SEQUENCES: 74

CORRESPONDENCE ADDRESS:

ADDRESSEE: Wolf, Greenfield & Sacks, P.C.

STREET: 600 Atlantic Avenue

CITY: Boston

STATE: MA

COUNTRY: U.S.A.

ZIP: 02210-2211

**COMPUTER READABLE FORM:**

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/978,404B

FILING DATE: 25-NOV-97

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/032,354

FILING DATE: 04-DEC-1996

ATTORNEY/AGENT INFORMATION:

NAME: Plumer, Elizabeth R.

REGISTRATION NUMBER: 35,637

REFERENCE/DOCKET NUMBER: B0801/7090

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-720-3500

TELEX:

TELEFAX: 617-720-2441

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 1128 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-978-404B-15

**RESULT 4**

US-09-016-366A-18

Sequence 18, Application US/09016366A

Patent No. 595431

**GENERAL INFORMATION:**

APPLICANT: Stevens, Richard L.

TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE

TITLE OF INVENTION: INHIBITORS

NUMBER OF SEQUENCES: 65

CORRESPONDENCE ADDRESS:

ADDRESSEE: Wolf, Greenfield & Sacks, P.C.

STREET: 600 Atlantic Avenue

CITY: Boston

STATE: MA

COUNTRY: U.S.A.

ZIP: 02210-2211

**COMPUTER READABLE FORM:**

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/016, 366A

FILING DATE: January 30, 1998

CLASSIFICATION: 550

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/037, 090

Qy 557 TGTGTCGCCGAGAACACCCGGAGGGACTCATGCCAGGGCAGTCCGGGAGGCCCCCTCGAGGGCAAC 198

Db 627 TGTGTCGCCGAGAACACCCGGAGGGACTCATGCCAGGGCAGTCCGGGAGGCCCCCTCGAGGGCAAC 137

Qy 617 GCAAGGTGAATGCCATCTACACCCGTCACCTACTACTTGACTGGATCCACCACT 736

Db 687 GCAAGGTGAATGCCATCTACACCCGTCACCTACTACAGGACAGCTGGTGCAGAGGATCATGTC 676

Qy 747 AGCCCAACGGCTGCGATCTACACCCGTCACCTACTTGACTGGATCCACCACT 746

Qy 737 ATGTCCCAAAGCGGTGAGCGGCC 733

Db 807 ATGTCCCAAAGCGGTGAGTCAGGC 833

Qy 257 ACCACAGTCTACACGCCAGATGGCTGGAGCGACTCGCCCTGCTGGAGGAGC 311

Db 327 ACCACAGTCTACACGCCAGATGGAGCGGACATGCCCTGCTGGAGGAGC 381

Qy 317 CGGTGAGGTCTCACGCCACCTCCAGACGGTACCCGGTACCTGGCCCTGCTCAGAGACCTCC 371

Db 387 CGGTGAGGTCTCACGCCAGTCAGGACAGCTGCTGGGTACCCGGTACCTGGCCCTGCTCAGAGACCTCC 441

Qy 437 CACCGCATTTCTCTGAGAGGTCAGGTCATGGCCATATAATGGAACACATTTGAGC 491

Db 507 CACCGCATTTCTCTGAGAGGTCAGGTCATGGCCATATAATGGAACACATTTGAGC 561

Qy 497 CAAATACACCTGGCCCTACACGGAGACGACCTGGCATGGCTGGAGCATGC 555

Db 567 CAAATACACCTGGCCCTACACGGAGACGACCTGGCATGGCTGGAGCATGC 628

Qy 557 TGTGTCGCCGAGACACCCGGAGGACTATGCCAGGGACTCCGGAGGGCCCTGGT 611

Db 627 TGTGTCGCCGAGACACCCGGAGGACTATGCCAGGGACTCCGGAGGGCCCTGGT 682

Qy 617 GCAAGGTGAATGCCACTGGCTGCGATCACCGGAGACGACCTGGCATGGCTGGCC 67

Db 687 GCAAGGTGAATGCCACTGGCTGCGATCACCGGAGACGACCTGGCATGGCC 74

Qy 677 AGCCCAACGGCTGCGATCTACACCCGTCACCTACTACTTGACTGGATCCACCACT 73

Db 747 AGCCCAACGGCTGCGATCTACACCCGTCACCTACTTGACTGGATCCACCACT 806

Qy 737 ATGTCCCAAAGCGGTGAGCGGCC 733

Db 807 ATGTCCCAAAGCGGTGAGTCAGGC 833

FILING DATE: 05-FEB-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Plumer, Elizabeth R.  
REGISTRATION NUMBER: 36,637  
REFERENCE/DOCKET NUMBER: B0801/7093  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-720-3500  
TELEFAX: 617-720-2441  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1137 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-09-016-366A-18

Query Match 94.8%; Score 731; DB 2; Length 1137;  
Best Local Similarity 98.7%; Pred. No. 6e-157; Matches 737; Conservatve 0; Mismatches 10; Indels 0; Gaps 0;

Qy 17 GAATGCGTCGGGGTAGGGAGGCCAGAGCAAGTGGCCCTGGAGGTGAGCTGAG 76  
Db 85 GCATGGTCGGGGTAGGGAGGCCAGAGCAAGTGGCCCTGGAGGTGAGCTGAG 144

Qy 77 TCCAGGCCATACTGGATGACTCTGGGGGGCTCCCTCATCCACCCCASTGGGTC 136  
Db 145 TCCAGGCCATACTGGATGACTCTGGGGGGCTCCCTCATCCACCCCASTGGGTC 204

Qy 137 TGACCGCCGGGGTAGGGAGGCCAGAGCAAGTGGCCCTCAGGATCTGGCCCCCTCAGGGTGAC 196  
Db 205 TGACGCCAGCAGCTGGACGGAGCTGGAGCTGGAGATCTGGCCCTCAGGGTGAC 264

Qy 197 TGCAGGAGGAGCACCTCTACTACAGGAGCAGCAGCTGCGGGTAGCAGAGATATCGTG 256  
Db 265 TGCAGGAGGAGCACCTCTACTACAGGAGCAGCAGCTGCGGGTAGCAGAGATATCGTG 324

Qy 257 ACCCACAGTCTACACCCGGCCAGATGGAGGGACATGCCCTGCTGGAGCTGGAGC 316  
Db 325 ACCCACAGTCTACACCCGGCCAGATGGAGGGACATGCCCTGCTGGAGCTGGAGGAGC 384

Qy 317 CGGAGGAGGCTCCAGGCCAACACSGTCACCCCTGGCCCTAGAGACCTTC 376  
Db 385 CGGAGGAGGCTCCAGGCCAACACSGTCACCCCTGGCCCTAGAGACCTTC 444

Qy 377 CCCCGGGATGCCGTTGGTCACTGGCTGGGGATGGAGCATGATGAGGCCCTTC 436  
Db 445 CCCCGGGATGCCGTTGGTCACTGGCTGGGGATGGAGCATGATGAGGCCCTTC 504

Qy 437 CACCCCATTCCTCTGAAAGCAGGTGAAGSTCCCTATAATGAAACCACATTGGAC 496  
Db 505 CACCCCATTCCTGAAAGCAGGTGAAGTCCCATAATGAAACCACATTGGAC 564

Qy 497 CAAATTACACCTTGCGCTACAGGGAGCAGGAGTCGGCATGGCCCTGAGCACATG 556  
Db 565 CAAATTACACCTTGCGCTACAGGGAGCAGGAGTCGGCATGGCCCTGAGCACATG 624

Qy 557 TGTGCGGGAGACCCGGAGGGACTCATGCCAGGGGACTCCGGAGGGCCCTGGT 616  
Db 625 TGTGCGGGAGACCCGGAGGGACTCATGCCAGGGGACTCCGGAGGGCCCTGGT 684

Qy 617 GCAAGGTGATGGCACTGGCTGCAAGGAGGGCTGGCTGAGCTGGGGAGGGCTGGCC 676  
Db 685 GCAAGGTGATGGCACTGGCTGCAAGGAGGGCTGGCCAGCTGGGGAGGGCTGGCC 744

Qy 677 AGCCAAACGCCCTGCACTACACCCGTCACCTACTTGACTGATGCCACT 736  
Db 745 AGCCAAACGCCCTGCACTACACCCGTCACCTACTTGACTGATGCCACT 804

Qy 737 ATGTCCTTAAAGCGTGAAGGCC 763  
Db 805 ATGTCCTTAAAGCGTGAAGGCC 831

Query Match 94.8%; Score 731; DB 2; Length 1137;  
Best Local Similarity 98.7%; Pred. No. 6e-157; Matches 737; Conservatve 0; Mismatches 10; Indels 0; Gaps 0;

Qy 17 GAATGCGTCGGGGTAGGGAGGCCAGAGCAAGTGGCCCTGGAGGTGAGCTGAG 76  
Db 85 GCATGGTCGGGGTAGGGAGGCCAGAGCAAGTGGCCCTGGAGGTGAGCTGAG 144

Qy 77 TCCAGGCCATACTGGATGACTCTGGGGGGCTCCCTCATCCACCCCASTGGGTC 136  
Db 145 TCCAGGCCATACTGGATGACTCTGGGGGGCTCCCTCATCCACCCCASTGGGTC 204

Qy 137 TGACCGCCGGGGTAGGGAGGCCAGAGCAAGTGGCCCTCAGGATCTGGCCCCCTCAGGGTGAC 196  
Db 205 TGACGCCAGCAGCTGGACGGAGCATGCCCTGCTGGGGTAGCAGAGATCTGG 264

Qy 197 TGCAGGAGGAGCACCTCTACTACAGGAGCAGCAGCTGCGGGTAGCAGAGATCTGG 316  
Db 325 ACCCACAGTCTACACCCGGCCAGATGGAGGGACATGCCCTGCTGGAGCTGGAGGAGC 384

Qy 317 CGGAGGAGGCTCCAGGCCAACACSGTCACCCCTGGCCCTAGAGACCTTC 376  
Db 385 CGGAGGAGGCTCCAGGCCAACACSGTCACCCCTGGCCCTAGAGACCTTC 444

QY 377 CCCGGGGATGCCGTCTGGTCAGGGCTGGGATGGAACATGATGAGGCCTCC 436  
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 445 CCCGGGGATGCCGTCTGGTCAGGGCTGGGATGGAACATGATGAGGCCTCC 504  
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 437 CACGCCCATTCCTCTGAGCAGGTGAAGTGTCCCATAATTGAAACACATTGTGAGC 496  
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 505 CACGCCCATTCCTCTGAGCAGGTGAAGTGTCCCATAATTGAAACACATTGTGAGC 564  
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 497 CAATATACCACCTGGCCTACAGGGAGACAGCTGGCAATGTCGCACTGGCACATGC 556  
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 565 CAATATACCACCTGGCCTACAGGGAGACAGCTGGCAATGTCGCACTGGCACATGC 624  
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 557 TGTGTGGGGACACGGGACTATGCCAGGGACTCGGAGGGACTCCGGGCTCTGGT 616  
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 625 TGTGTGGGGACACGGGACTATGCCAGGGACTCGGAGGGACTCCGGGCTCTGGT 684  
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 617 GCAAGGTGAATGGCACCTGGCTGAGGGGGGNGTCACTGGCTGGGAGGGCTGG 676  
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 685 GCAAGGTGAATGGCACCTGGCTGAGGGGGGNGTCACTGGCTGGGAGGGCTGG 744  
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 677 AGCCCAACGGCTGCACTACACCCGAGTCACCTACTACTTGACTTGATCACC 736  
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 745 AGCCCAACGGCTGCACTACACCCGAGTCACCTACTACTTGACTTGATCACC 804  
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 737 ATGCCCAAAAGCCGTGAAGGCC 763  
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 805 ATGCCCAAAAGCCGTGAAGGCC 831  
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 RESULT 6  
 US-09-970A-1  
 ; Sequence 1, Application US/09079970A  
 ; General Information:  
 ;   APPLICANT: Maffitt, Mark A.  
 ;   APPLICANT: Haak Frendscho, Mary  
 ;   TITLE OF INVENTION: Enzymatically-Active Recombinant Human  
 ;   NUMBER OF SEQUENCES: 6  
 ;   CORRESPONDENCE ADDRESS:  
 ;   ADDRESSEE: Intellectual Property Department  
 ;   STREET: 8000 Excelsior Drive, Suite 401  
 ;   CITY: Madison  
 ;   STATE: WISCONSIN  
 ;   COUNTRY: U.S.A.  
 ;   ZIP: 53717-1914  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/079,970A  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ;   NAME: Leone, Joseph T.  
 ;   REGISTRATION NUMBER: 37,170  
 ;   REFERENCE/DOCKET NUMBER: 34506.073  
 ; TELEPHONE: (608) 831-2106  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ;   SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 735 base Pairs  
 ;   TYPE: nucleic acid  
 ;   STRANDEDNESS: single  
 ;   TOPOLOGY: linear  
 ;   MOLECULE TYPE: DNA (genomic)  
 ;   HYPOTHETICAL: NO  
 ; ANTI-SENSE: NO  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 1..735  
 ; US-09-970A-1  
 ; Query Match 94.5%; Score 728.6; DB 4; Length 735;  
 ; Best Local Similarity 99.5%; Pred. No. 1..9e-156;  
 ; Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 ; 1 ATCGTGTGGGGTCAGGAGGCCAGAGCACTGGCCCTGAGAGTC 78  
 ; Db 1 ATCGTGTGGGGTCAGGAGGCCAGAGCACTGGCCCTGAGAGTC 60  
 ; Db 79 CACGCCCATACTGTGACTTCGGGGCTCCATCCCCAGTGGTCTG 138  
 ; Db 61 CACGCCCATACTGTGACTTCGGGGCTCCATCCCCAGTGGTCTG 120  
 ; QY 139 ACCGGCGGGGGTGGGGAGCCGGACGGTCAAGGATCTGGCCCTAGGGTCAACTG 198  
 ; Db 121 ACCGAGCGCCTGCGGGGAGCCGGACGGTCAAGGATCTGGCCCTAGGGTCAACTG 180  
 ; QY 199 CGGGAGCAGCACCTCTACTACCGACGCTGTCGAGGTCACTGTGAC 238  
 ; Db 181 CGGGAGCAGCACCTCTACTACCGACGCTGTCGAGGTCACTGTGAC 240  
 ; Qy 259 CACAGTTCAACCCCCAGTGGAGCGACATGCCCTCTGGAGCTGGAGGCC 318  
 ; Db 241 CCACAGTTCAACGCCAGTGGAGCGACATGCCCTCTGGAGCTGGAGGCC 300  
 ; QY 319 GTGAAGGTCTCACGGTCACACGGTCACCGGTACCCCTGCCCTCAGAACCTTCCC 378  
 ; Db 301 GTGAAGGTCTCACGGTCACACGGTCACCGGTACCCCTGCCCTCAGAACCTTCCC 360  
 ; QY 379 CCGGGATCGCTGCTGGTCACTGGCTGGCCATGTCGACATGAGCGCTCCCA 438  
 ; Db 361 CCGGGATCGCTGCTGGTCACTGGCTGGCCATGTCGACATGAGCGCTCCCA 420  
 ; QY 439 CGGGCATTTCTCTGAGCAGGAGGTGAGGTCCCTAATGAAACACATTGTGACCA 498  
 ; Db 481 AAATACCACTTGGCTTACACGGAGACGACCTGGCATGTCGCGTGAACATGTC 540  
 ; QY 421 CCGGGATTCCTCTGAGCAGGAGGTGAGGTCCCTAATGAAACACATTGTGACCA 480  
 ; Db 559 TGTGGGGACACCGGGAGGAGCTACATGCCAGGGACTCTCGGAGGGCCCTCTGGT 618  
 ; Db 541 TGTGGGGACACCGGGAGGAGCTACATGCCAGGGACTCTCGGAGGGCCCTCTGGT 600  
 ; QY 619 AGGGTAATGGCAGCTGGCTGAGGGGGCTGTCACCTGGGGAGGGCTGRCGAG 678  
 ; Db 601 AGGGTAATGGCAGCTGGCTGAGGGGGCTGTCACCTGGGGAGGGCTGRCGAG 660  
 ; QY 679 CCCAACCGGCTGGCATCTACCCGGTGTCACTACTTGACTGGATCACCCTAT 738  
 ; Db 661 CCCAACCGGCTGGCATCTACCCGGTGTCACTACTTGACTGGATCACCCTAT 720  
 ; QY 739 GTCCCCAAAAGCCG 753  
 ; Db 721 GTCCCCAAAAGCCG 735  
 ; RESULT 7  
 ; US-09-016-366A-22  
 ; Sequence 22, Application US/09016366A  
 ; General Information:  
 ;   APPLICANT: Stevens, Richard L.  
 ;   APPLICANT: Huang, Chifu  
 ;   TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE  
 ;   TITLE OF INVENTION: INHIBITORS

NUMBER OF SEQUENCES: 65  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
 STREET: 600 Atlantic Avenue  
 CITY: Boston  
 STATE: MA  
 COUNTRY: U.S.A.  
 ZIP: 02210-2211  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSEQ for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/016,366A  
 FILING DATE: January 30, 1998  
 CLASSIFICATION: 530  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 60/037,090  
 FILING DATE: 05-FEB-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Plumer, Elizabeth R.  
 REGISTRATION NUMBER: 36,637  
 REFERENCE/DOCKET NUMBER: B0801/7093  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-720-3500  
 TELEX: 617-720-2441  
 INFORMATION FOR SEQ ID NO: 22:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1081 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 US-00-016-366A-22

Query Match 94.2%; Score 726.2; DB 2; Length 1081;  
 Best Local Similarity 98.3%; Pred. No. 7.2e-156;  
 Matches 734; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 17 GAATGCTGGGGTAGGAGGCCAGGACAAGTGGCCCTGGAGGTGAGAG 76  
 Db 66 GCATCGTTGGGGTAGGAGGCCAGGACAAGTGGCCCTGGAGGTGAGAG 125  
 QY 77 TCCAGGCCATACTGGACTCTGGGGGCTCCATCCACCCCASTGGGGC 136  
 Db 126 TCCGGACCCATACCTGGACTCTGGGGGCTCCATCCACCCCASTGGGTGC 185  
 QY 137 TGACGCCGCGGGCTGGAGCCGCTCAAGATCTGGCCCTCAGGGCAC 196  
 Db 186 TGACGCCGCGACTCTGGGGGCTGGAGCTGAAGATCTGGCCCTCAGGGCAC 245  
 QY 197 TCGGGAGCAGCACCTCTACTACAGGACAGCTGGGGTGAAGCAGGATCATCGTC 256  
 Db 246 TCGGGAGCAGCACCTCTACTACAGGACAGCTGGGGTGAAGCAGGATCATCGTC 305  
 QY 257 ACCCACTGCTACACGCCAGATGGAGGGAGATGCCCTGGAGCTGGAGAC 316  
 Db 306 ACCCACTGCTACACGCCAGATGGAGGGAGATGCCCTGGAGCTGGAGAC 365  
 QY 317 CGCGAGGCTCCGCCAACGTCACCGTCACCCCTGGCTAGAGCTTC 376  
 Db 366 CGGAAAGGTCTCCAGCCACGGTCCACCGGTCTGGAGACTTC 425  
 QY 377 CCCCGGGATGCTCCGCCAACGTCACCGTCACCCCTGGCTAGAGCTTC 436  
 Db 426 CCCCGGGATGCTCCGCCAACGTCACCGTCACCCCTGGCTAGAGCTTC 485  
 QY 437 CACGCCATTCCTCTGAAGCAGGGTAAGSTGCCCTAATGAAACCATGTTGAG 496  
 Db 486 CACCCCATTCCTCTGAACGGGAGCTCCCTAATGAAACCATGTTGAGC 545

RESULT 8  
 US-00-978-404B-17  
 ; Sequence 17, Application US/08978404B  
 ; Patent No. 5968782  
 GENERAL INFORMATION:  
 APPLICANT: Stevens, Richard L.  
 TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES  
 TITLE OF INVENTION: FIBRINOLYNE  
 NUMBER OF SEQUENCES: 74  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
 STREET: 600 Atlantic Avenue  
 CITY: Boston  
 STATE: MA  
 COUNTRY: U.S.A.  
 ZIP: 02210-2211  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSEQ for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/978, 404B  
 FILING DATE: 25-NOV-97  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 60/032, 354  
 FILING DATE: 04-DEC-1996  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Plumer, Elizabeth R.  
 REGISTRATION NUMBER: 36,637  
 REFERENCE/DOCKET NUMBER: B0801/7090  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-720-3500  
 TELEX: 617-720-2441  
 INFORMATION FOR SEQ ID NO: 17:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1081 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 US-00-978-404B-17

Query Match 94.2%; Score 726.2; DB 2; Length 1081;  
 Best Local Similarity 98.3%; Pred. No. 7.2e-156;  
 Matches 734; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 17 GAATGCTGGGGTAGGAGGCCAGGACAAGTGGCCCTGGAGGTGAGAG 76  
 Db 66 GCAAGGTGATGGCACCTGGCTGAGGGGAGACGAGTCGGAGGACATGC 605  
 Db 606 TGTGCGGGACACCCGGAGGGACTCATGCCAGGGGACTCGGAGGGCCTGGGT 616  
 QY 617 GCAAGGTGATGGCACCTGGCTGAGGGGAGACGAGTCGGAGGACATGC 665  
 Db 666 GCAAGGTGATGGCACCTGGCTGAGGGGAGCGGGCTGGTGCAGTGGGGCTGGCC 725  
 QY 677 AGCCCAACGGCCUGGCCAGTACACCGTGACTACTTGACTCTGGACATCCACCACT 736  
 Db 726 AGCCCAACGGCCUGGCCAGTACACCGTGACTACTTGACTCTGGACATCCACCACT 785  
 QY 737 ATGTCCCCCAAAAGCCGTGAGTCAGGC 763  
 Db 786 ATGTCCCCCAAAAGCCGTGAGTCAGGC 812

QY 497 CAAATACACCTGGCCCTACGGAGAGCTGCCATGTCCTGGAGGACATGC 556  
 Db 546 CAAATACACCTGGCCCTACGGAGAGCTGCCATGTCCTGGAGGACATGC 605  
 QY 557 TGTGCGGGACACCCGGAGGGACTATGCCAGGGACTCGGAGGGCCTGGGT 616  
 Db 606 TGTGCGGGACACCCGGAGGGACTATGCCAGGGACTCGGAGGGCCTGGGT 665  
 QY 617 GCAAGGTGATGGCACCTGGCTGAGGGGAGACGAGTCGGAGGACATGC 676  
 Db 666 GCAAGGTGATGGCACCTGGCTGAGGGGAGCGGGCTGGTGCAGTGGGGCTGGCC 725  
 QY 677 AGCCCAACGGCCUGGCCAGTACACCGTGACTACTTGACTCTGGACATCCACCACT 736  
 Db 726 AGCCCAACGGCCUGGCCAGTACACCGTGACTACTTGACTCTGGACATCCACCACT 785  
 QY 737 ATGTCCCCCAAAAGCCGTGAGTCAGGC 763  
 Db 786 ATGTCCCCCAAAAGCCGTGAGTCAGGC 812

77 TCCACGGCCATACTGGATGCACTTGTGCGGGCTCCTCATCACCCCAGTGGGGC 136  
 Db TCGGGACGCCATACTGGATGCACTTGTGCGGGCTCCTCATCACCCCAGTGGGC 185  
 126 TGACGCCGCGGGCGTGGACGGGCGACCTCAAGATCTGGCCGCGCTTAAGGGTCAC 196  
 Db 186 TGACGCCGAGGCCACTCTGGTGCGACCGGACCTCAAGATCTGGCCGCGCTTAAGGGTCAC 245  
 Qy 197 TCGGGAGCGCACCTCTACTACCGAGGACAGTGTGCGGCTCACAGGATCATGTG 256  
 Db 246 TCGGGAGCACGACCTCTACTACCGAGGACAGTGTGCGGCTCACAGGATCATGTG 305  
 Db 257 ACCCACAGTTCTACACGCCAGATGGAGGAGATCGCCCTGCTGGAGGAGC 316  
 Db 306 ACCCACAGTTCTACACGCCAGATGGAGGAGATCGCCCTGCTGGAGGAGC 365  
 Qy 317 CGGTGAGGCTCCAGGCCAGTCCACGCCAGTCCACGCCCTGCTGGAGGAGC 376  
 Db 366 CGGTGAGGCTCCAGGCCAGTCCACGCCAGTCCACGCCCTGCTGGAGGAGC 425  
 Qy 377 CCCGGGATGCCGCTGTGCTACTGGCTGGGCGATGGACATGAGGCCCTC 436  
 Db 426 CCCCGGGATGCCGCTGTGCTACTGGCTGGGCGATGGACATGAGGCCCTC 485  
 Qy 437 CACCGCCATTCTCTGTGAAAGGAGGTGAGGAGCTCCATAATGAAACCACTTGTGAC 496  
 Db 486 CACCGCCATTCTCTGTGAAAGGAGGTGAGGAGCTCCATAATGAAACCACTTGTGAC 545  
 Qy 497 CAAATACTACCTTGCGCTTACAGGGAGACGAGCTCCATGRCGCGACATG 556  
 Db 546 CAAATACTACCTTGCGCTTACAGGGAGACGAGCTCCATGRCGCGACATG 605  
 Qy 557 TGTGCGCCGACACCCGGAGGACTCAAGGGCGACTCCAGGGCGCTGGTCACTGGGGCCCTGGT 616  
 Db 606 TGTGCGCCGACACCCGGAGGACTCAAGGGCGCTGGTCACTGGGGCCCTGGT 665  
 Qy 617 GCAAGGTGAATGGCACCTGTGTGAGGGCGCTGGGCGAGGACTGTGCCCC 676  
 Db 666 GCAAGGTGAATGGCACCTGTGTGAGGGCGCTGGGCGAGGACTGTGCCCC 725  
 Qy 677 AGCCCAACCCGCGCATCTACACCGTGTGACCTACTACTTGACTGTGATCCACT 736  
 Db 726 AGCCCAACCCGCGCATCTACACCGTGTGACCTACTACTTGACTGTGATCCACT 785  
 Qy 737 ATGCCCCAACAGCGTGAAGGGCC 763  
 Db 786 ATGCCCCAACAGCGTGAAGGGCC 812

RESULT 9

US-09-016-366A-16

Sequence 16, Application US/09016366A

Patent No. 5955431

GENERAL INFORMATION:

APPLICANT: Stevens, Richard L.

APPLICANT: Huang, Chifu

TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE

TITLE OF INVENTION: INHIBITORS

NUMBER OF SEQUENCES: 65

CORRESPONDENCE ADDRESS:

ADRESSEE: Wolf, Greenfield & Sacks, P.C.

STREET: 600 Atlantic Avenue

CITY: Boston

STATE: MA

COUNTRY: U.S.A.

ZIP: 02210-2211

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/016,366A

FILING DATE: January 30, 1998

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/037,090

FILING DATE: 05-FEB-1997

ATTORNEY/AGENT INFORMATION:

NAME: Plumer, Elizabeth R.

REGISTRATION NUMBER: 36, 637

REFERENCE/DOCKET NUMBER: B0801/7093

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-720-3500

TELEFAX: 617-720-2441

INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:

LENGTH: 1154 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-09-016-366A-16

Query Match 89.0%; Score 686.2; DB 2; Length 1154;

Best Local Similarity 94.3%; Pred. No. 8.2e-147;

Matches 712; Conservative 0; Mismatches 43; Indels 0; Gaps 0

Qy 17 GAATCGCGGGTCAAGGGCCCGAGGAGCAGTGGCCCTGGAG 76

Db 106 GTATCGTGGGGTCAGGGGCCAGAGCAAGTGGCCCTGGAG 165

Qy 77 TCCAGGCCCATACTGGTACCTTCTGGGGGCTCCATCACCCACAGTGGTC 136

Db 166 TCCGGACCCATACTGGTACCTTCTGGGGGCTCCATCACCCACAGTGGTC 225

Qy 137 TGACCGCGGGTGTGAGGAGGACTCAAGGAGCTGGCGGCCAGGGTCAC 196

Db 226 TGACCGCGGGTGTGAGGAGGACTCAAGGAGCTGGCGGCCAGGGTCAC 285

Qy 197 TCGGGAGGACCTCTACTACCGGACAGCTGTGCGGTCAGCAGGATCATGTG 256

Db 286 TCGGGAGGACCTCTACTACCGGACAGCTGTGCGGTCAGCAGGATCATGTG 345

Qy 257 ACCCACAGTTCTACACCCGCCAGATGGAGGGAGCTGGAGCTGGAGGAC 316

Db 346 ACCCACAGTTCTACATCCAGACTGGAGGGAGATGGGGCTGGAGCTGGAGGAC 405

Qy 317 CGGTGAGGCTCCAGGACCTGAGGAGCTGGAGGGAGCTGGAGCTGGAGGAC 376

Db 406 CGGTGAGGCTCCAGGACCTGAGGAGCTGGAGGGAGCTGGAGCTGGAGGAC 465

Qy 377 CCCGGGATGCCGCTGTGGTACTGGCTGGGCGATGTGACATGAGGCCCTC 436

Db 466 CCCGGGATGCCGCTGTGGTACTGGCTGGGCGATGTGACATGAGGCCCTC 525

Qy 437 CACGGCCATTCTCTGTGAAAGGAGTGTGACCTGGCCCTAATGGAAACCACTTGTGAC 496

Db 526 CACGGCCATTCTCTGTGAAAGGAGTGTGACCTGGCCCTAATGGAAACCACTTGTGAC 585

Qy 497 CAAATACTACCTTGCGCTTACAGGGAGACGAGCTGGCAAGGGAGCTGGAG 556

Db 586 CAAATACTACCTTGCGCTTACAGGGAGACGAGCTGGCAAGGGAGCTGGAG 645

Qy 617 GCAAGGTGAATGGCACCTGTGCTGGGGCTGGAGCTGGAGGGAGCTGGAG 676

Db 706 GCAGGGTGAATGGCACCTGTGCTGGGGCTGGAGCTGGAGGGAGCTGGAG 765

Qy 677 AGCCCAACCGGCCCTGCACTACACCGCGTCACTACTTGACTGTGATCCACT 736

Db 766 AGCCCAACCGGCCCTGCACTACACCGCGTCACTACTTGACTGTGATCCACT 825

RESULT 10  
; Sequence 11, Application US/08978404B  
; Patent No. 5968782  
; GENERAL INFORMATION:  
; APPLICANT: Stevens, Richard L.  
; TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES  
; TITLE OF INVENTION: FIBRINOGEN  
; NUMBER OF SEQUENCES: 74  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
; STREET: 600 Atlantic Avenue  
; CITY: Boston  
; STATE: MA  
; COUNTRY: U.S.A.  
; ZIP: 02210-2211  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/978,404B  
; FILING DATE: 25-NOV-97  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/032,354  
; FILING DATE: 04-DEC-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Plumer, Elizabeth R.  
; REGISTRATION NUMBER: 36,637  
; REFERENCE/DOCKET NUMBER: B0801/7090  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-720-3500  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1154 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-08-978-404B-11  
; Query Match 89.0%; Score 686.2; DB 2; Length 1154;  
; Best Local Similarity 94.3%; Pred. No. 8.2e-147;  
; Matches 712; Conservative 0; Mismatches 43; Indels 0; Gaps 0;  
; QY 17 GAATGCTGGGGTCAGGAGGCCAGGGCAAGTGGCCCTGAGGAGAG 76  
; Db 106 GTATGCTGGGGTCAGGAGGCCAGGGCAAGTGGCCCTGAGGAGAG 165  
; QY 77 TCCAGGGCCATACTGGACACTGCTGGGGGCTCCATCCACCCCACTGGGTC 136  
; Db 166 TCCGGGACCACTACTGGACACTGCTGGGGGCTCCATCCACCCCACTGGGTC 225  
; QY 197 TGCGGGAGCACCTCTACTACCAAGGACCAAGCAGCTGGGGTCAGCAAG 256  
; Db 137 TGACCCGGGGGGTGCGGGGACGCCAAGGATCTGGCCCTCAAGGGTCAAC 196  
; Db 226 TGACCGGGGGGCACTGCTGGGACGGGACTGCAAGGATCTGGCCCTCAAGGGTCAAC 285  
; Db 346 ACCCACAGTCTACACCGCCAGATGGAGGAGATCCCTCTGGAGGTGGAGGAC 405  
; QY 257 ACCCACAGTCTACACCGCCAGATGGAGGAGATCCCTCTGGAGGTGGAGGAC 316  
; Db 346 ACCCACAGTCTACACCGCCAGATGGAGGAGATCCCTCTGGAGGTGGAGGAC 405

RESULT 11  
; Sequence 7, Application US/08978404B  
; Patent No. 5968782  
; GENERAL INFORMATION:  
; APPLICANT: Stevens, Richard L.  
; TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES  
; TITLE OF INVENTION: FIBRINOGEN  
; NUMBER OF SEQUENCES: 74  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
; STREET: 600 Atlantic Avenue  
; CITY: Boston  
; STATE: MA  
; COUNTRY: U.S.A.  
; ZIP: 02210-2211  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/978,404B  
; FILING DATE: 25-NOV-97  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/032,354  
; FILING DATE: 04-DEC-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Plumer, Elizabeth R.  
; REGISTRATION NUMBER: 36,637  
; REFERENCE/DOCKET NUMBER: B0801/7090  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-720-3500  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1219 base pairs



Db 605 GGAAGTACCACTGGCCTCATACAGGGAGATGATTGCCATGTGCCATGCGATGC 664  
 Qy 557 TGTGTGCCGGAACACCCGGAGGACTCATGCCAGGGACTCGCAGGGCTCAGGGTGC 616  
 Db 665 TGTGCTGAAATACCAAGAGAGACTCTGCCAGGGATTCAGGGGCCACTGGCT 724  
 Qy 617 GCAAGGTGAATGCCACTTGCTCAGGGGGCTGTCAGCTGGGGACTCGGCC 676  
 Db 725 GCAAGTGAAGGTACCTGGCTCAGGGAGCTGGTCACTGGGGCTGCGAC 784  
 Qy 677 AGCCAAACGGCTGGCATCTACACCCGGTCACTACTACTTGAGCTGATCCACCT 736  
 Db 785 AGCCAAACAGCTGGCATCTACACCCGGTCACTACTACTTGAGCTGATCCACCT 844  
 Qy 737 ATGCCAACAGCTGGCATCTACACCCGGTCACTACTACTTGAGCTGATCCACCT 864  
 Db 845 ATGCCCTGAGCATCTCTGA 864

RESULT 13  
 US-08-978-404B-20  
 ; Sequence 20 Application US/08978404B  
 ; Patent No. 596872  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Stevens, Richard L.  
 ; TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES  
 ; TITLE OF INVENTION: FIBRINOLYtic  
 ; NUMBER OF SEQUENCES: 74  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
 ; STREET: 600 Atlantic Avenue  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: U.S.A.  
 ; ZIP: 02210-2211  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastS2Q for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/978,404B  
 ; FILING DATE: 25-NOV-97  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 60/032,354  
 ; FILING DATE: 04-DEC-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Plumer, Elizabeth R.  
 ; REGISTRATION NUMBER: 36,637  
 ; REFERENCE/DOCKET NUMBER: B0801/7090  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 617-720-3500  
 ; TELEFAX: 617-720-2441  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 20:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1108 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; US-08-978-404B-20

Query Match 60.9%; Score 469.6; DB 2; Length 1108;  
 Best Local Similarity 77.2%; Pred No. 8.1e-98; Matches 571; Conservative 0; Mismatches 169; Indels 0; Gaps 0;

Qy 17 GAATGCTGGGGCTCAGGAGGCCACAGGCAAGTGGCCCTGAGCTGAGAG 76  
 Db 125 GCAATGCTGGAGGACATGAGCTTCTGAGAGTAATGAGCTGGAGGCTGAGAT 184  
 Qy 77 TCCAGGCCATACTGGCACTCTGGGGCTCCATCACCCCCAGTGGTGC 136

Db 185 TAAATTAACATCTGGATACATTCTCGGGGCTCTCATCCACCAAGTGGTGC 244  
 Qy 137 TGAACGCCCGCTGCTGGACGGAGCTGCAAGGACTCTGGCCCTCAAGGTGAC 196  
 Db 245 TACTGCGCAACTGTGTTGGGACGACATCAAAGCCACAGCTCTCCGGTGCAGC 304  
 Qy 197 TGGGGAGCAGCACCTACTACAGGACAGCTGCTGGCAGGAGCATCTGGTGC 256  
 Db 305 TCTGTGASCACTATCTACTATGGGACCAAGCTCTCTTGAACGGATCTGGTGC 364  
 Qy 257 ACCACAGTCTACACGCCAGATGGAGGGACATGCCCTGCTGGAGCTGGAGGC 316  
 Db 365 ACCCCACTATACACGCCAGGGGGAGGGAGCTGGCCCTGCTGGAGCTGGTGC 424  
 Qy 317 CGGGAASGTCCTCAGCCAGTGCACAGGTACCGTACCGCCCTGCTGGAGCTTC 376  
 Db 425 CTGTGAATGTCCTCACCATATCCACCCATATCCCTGCCCTGCTGGAGACCTTC 484  
 Qy 377 CCCGGGGATGCCGTGGTCACTGCCTGGCGCATGTGACAATGTGAGGCGCTCC 436  
 Db 485 CCCCTGGGACATCCTGGCTGGTGAAGGGCATGGTCAAGGACATGGGCTTC 544  
 Qy 437 CACGCCATTTCTCTGAAGCAGTGAAGGTCATGGAAACACATTTGTGAC 496  
 Db 545 CACCTCTATCTCTGAAGCAGTGAAGGTCATGGAAACACATTTGTGAC 604  
 Db 497 CAAATACACCTTGGCCTCATACGGGAGACGAGTGGCAGCTGGCCCTGCTGGAC 556  
 Db 605 GAAAGTACCACTGGCCTCATACAGGGAGATGATTCCATGTGCAATGAGC 664  
 Qy 557 TGTGTGCCGGAACCCGGAGGAGCTATGCCAGGGACTCTGGAGGGACTCTGG 616  
 Db 665 TGTGCTGAAATACCAAGAGAGACTCTGCCAGGGGACTCTGGGGCTACTGGT 724  
 Qy 617 GCAAGGTGAATGCCACTTGCTCAGGGAGATGATTCCATGTGCAATGAGC 676  
 Db 725 GCAAGTGAAGGTACCTGGCTCAGGGAGGACTGGTCACTGGGCTGCCC 784  
 Qy 677 AGCCAAACGGCCCTGGCATCTACACCCGGTCACTACTACTTGAGCTGATCCACCT 736  
 Db 785 AGCCAAACAGCTGGCATCTACACCCGGTCACTACTACTTGAGCTGATCCACCT 844  
 Qy 737 ATGCCAACAGCTGGCATCTACACCCGGTCACTACTACTTGAGCTGATCCACCT 864  
 Db 845 ATGCCCTGAGCATCTCTGA 864

RESULT 14  
 US-08-978-404B-1  
 ; Sequence 1, Application US/08978404B  
 ; Patent No. 596872  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Stevens, Richard L.  
 ; TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES  
 ; TITLE OF INVENTION: FIBRINOLYtic  
 ; NUMBER OF SEQUENCES: 74  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
 ; STREET: 600 Atlantic Avenue  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: U.S.A.  
 ; ZIP: 02210-2211  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastS2Q for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/978,404B  
 ; FILING DATE: 25-NOV-97  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/032,354  
 FILING DATE: 04-DEC-1996  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Plumer, Elizabeth R.  
 REGISTRATION NUMBER: 35,637  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-720-3500  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1031 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 US-08-978-404B-1

Query Match 60.1%; Score 463; DB 2; Length 1031;

Best Local Similarity 76.4%; Pred. No. 2.5e-96; Mismatches 0; Gaps 0; Matches 568; Conservative 0; Indels 0; Gaps 0;

QY 14 AAGGAGATCGTCGGGGCTCAGGAGGCCAGGAGCACTGGAAAGTGCCCTGGCAGGTGGCTGA 73  
 106 AAGGCATGTTGGGGACAGGAGGCCAGGAGCACTGGAAAGTGCCCTGGCAGGTGGCTGA 165  
 Db 74 GAGTCACGCCCATACTGATGACTTCTGCGGGGRCCTCATCCACCCAGTGG 133  
 Qy 166 GTGCCATGACACTACTGGATGATCATTCTGGGGGGCTCATCCACCCAGTGG 225  
 Db 134 TGCCTGACCGCCGGCGCGTGGGACCGACGCCAGGATCTGGCCGCCCTCAGGGTC 193  
 Qy 226 TGCTRACTGGCACACTGTGGGACCGATTTGCTGAGTGGAGTAC 285  
 Db 194 AACTCGGGGAGCAGCACCTACTACAGGACCACTGTCGCGGTCAAGGAGTATCG 253  
 Qy 286 AGCUCGGTAAGCAATGACGATTCATGACACTGTGACTGAGCTGGACCGACATCA 345  
 Qy 254 TGCACCCACAGTCTCTACACGCCGAGATCGGAGGACATCGCCCTGCGAGCTGGAG 313  
 Db 346 CACACCCGACTCTACATGTCAGGATGGGGAGACATGCCCCTGCTCAGAAC 405  
 Qy 314 AGCCGGTGAGGTCTCCACGCCAACGTCACCGTACCTGCCCTGCTCAGAGACCT 373  
 Db 406 ACCCTGTGACATTTCTGACTATGTCACCTCTGCTCCTACCTCTGCTCAGAAC 465  
 Qy 374 TCCCCCGGGGATGGCTGCTGGGACTGCTGAGCAATGATGAGCGCC 433  
 Db 466 TCCCCCTCAGAACGTTGCTGGGATCACATGACATGGTAAACC 525  
 Qy 434 TCCCCCGGCCATTCTCTGAAAGGTTGAGGTCCTCCATAATGGAAACACATTGTTG 493  
 Db 526 TGCCCGCACCATTCCTTGAGGGAGGTGCAAGTCTCCATATAGAAACACCCTTGG 585  
 Qy 494 ACGCAAATACCTTGGCCATCACCGGAGACGCTCCGATCCTGGTGAAGGACA 553  
 Db 586 ACTCTGAGTACAAAGGCTCTACACAGGTGACATGTCAGGATGAGTGA 645  
 Qy 554 TGCCTGTCGGGACACCCGGAGACTCATGCCAGGGCAGTCCGGAGGCCCTGG 613  
 Db 646 TGCCTGTCGGGATGAGGAGCATGACTCTCCAGGGGACTCCGGAGGACTCTGG 705  
 Qy 614 TGTGGAGGGAATGCGACCTGGCTGCAGCGGGGCTCAGCTGGGAGCTGG 673  
 Db 706 TCTGCAAGGTTAGAACACACCTGGCTGCAAGGGGTGGTCAGCTGGGGTGG 765  
 Qy 674 CCCACCCAAACCGGCTGCACTACACCGTGTACCTACTGACTGGATCCAC 733  
 Db 766 CACACCCAAACGGCTGCACTACACCGGGTGGTCACCTATTACTGGACTGGATCCAC 825  
 Qy 734 ACTATGCCCCAAAGGCTGA 756  
 Db 826 ACTATGCCCCAAGGACTCTGGA 848

APPLICATION NUMBER: 60/032,354  
 FILING DATE: 04-DEC-1996  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Plumer, Elizabeth R.  
 REGISTRATION NUMBER: 35,637  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-720-2441  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1103 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 US-09-016-366A-24

RESULT 15

US-09-016-366A-24

Sequence 24, Application US/09016366A

Patent No. 5955431

GENERAL INFORMATION:

APPLICANT: Stevens, Richard L.

TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE

TITLE OF INVENTION: INHIBITORS

NUMBER OF SEQUENCES: 65

CORRESPONDENCE ADDRESS:

ADDRESSEE: Wolf, Greenfield & Sacks, P.C.

STREET: 600 Atlantic Avenue

CITY: Boston

STATE: MA

COUNTRY: U.S.A.

ZIP: 02210-2211

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/016,366A

FILING DATE: January 30, 1998

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/037,090

FILING DATE: 05-FEB-1997

ATTORNEY/AGENT INFORMATION:

NAME: Plumer, Elizabeth R.

REGISTRATION NUMBER: 35,637

REFERENCE/DOCKET NUMBER: B0801/7093

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-720-3500

TELEFAX: 617-720-2441

INFORMATION FOR SEQ ID NO: 24:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 1103 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA

US-09-016-366A-24

Query Match 57.8%; Score 445,6; DB 2; Length 1103;  
 Best Local Similarity 75.1%; Pred. No. 2.2e-92; Mismatches 184; Indels 0; Gaps 0;  
 Matches 556; Conservative 0; Mismatches 184; Indels 0; Gaps 0;

QY 17 GAATGCTGGGGTCAGGAGGCCAGGAGCAGTGGCCCTSGAGGAGCTGAG 76  
 102 GCATGTTGGGGATGAGGAGCTCTGTAAGTGTAGTGGCCCTGGCAGGAGCTGAG 161  
 Db 77 TCCAGGGCCATACCTGGATGCACTCTGCGGAGCTCCATCACCCCTGGGTC 136  
 Qy 162 TAAATTCACCTCTGGATGCACTCTGCGGCTCCCATTCACCCACAGTGGTC 221  
 Db 137 TAGCCCGGGCGCTGGTGGACGGACTCAAGATCTGGCCCTCAGGTGCAAC 96  
 Db 222 TCACTGGCCACACTGTGCGGAGTCACAGGCTGGTACAG 281  
 Qy 197 TGGGGAGCAGCACCTCTACTACAGGAGCAGTGGCTGGCAGCAGGATCATCG 256  
 Db 282 TTGGAGGAGTACTATGCTGGACCGACCTGTGACCGACCTTGTC 341  
 Qy 257 ACCCAGAGTCTACACGCCAGTGGAGGAGCTCCCTGTGGAGTGGAGGAGC 316  
 Db 342 ACCCCACTACTACACAGTGGAGGATGGGGAGACATGGCCCTGCTGGAGCTTGC 401  
 Qy 317 CGGTGAAGGCTCCAGCCACGCTCACCGTACCCCTGCCCCCTGCTCAGAGACCTTGC 376

Search completed: January 31, 2003, 05:25:27  
Job time : 72 secs



	/organism="Homo sapiens"
	/lab_xref="taxon;606"
	/clone="IMAGE:5748499"
	/lab_host="DH10B"
	/note="Organ: pooled pancreas and spleen; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of spleen and pancreas from 28 yo male. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-2.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 025. Note: this is a NIH_MGC Library."
BASE COUNT	180
ORIGIN	a
Query Match	87.7%; Score 676.2; DB 14; length 1030; Best Local Similarity 94.5%; Pred. No. 6.9e-133; Matches 711; Conservative 0; Mismatches 39; Indels 2; Gaps 1;
Qy	17 GAACTGCTGGGGCAGAGGCCACAGGACACTGGCCCTGGCAGGTGAGAG 76
Dy	121 GTATGCTCGGGGTTAGGAGGCCACAGGACACTGGCCCTGGCAGGTGAGAG 180
Qy	77 TCCACGGCCCATACTGGACACTCTGGGGGACTCCATCCACCCCCAGTGGGTGC 136
Dy	181 TCCGGGACCGATACTGGATGCACTTCCTGGGGCTCCATCCACCCCAGTGGGTGC 240
Qy	137 TGACCGCCGGGTGCGGACCCAGGAGCTGAGAAGTCAGAGATCTGGCAC 196
Dy	241 TGACCGCCGGCACTGCTGGGACCGTCAGATCTGGCACCCCTCAGGTGCAC 300
Qy	197 TGGGGGAGGCTCCAGCCCTCACTACAGGAGCAGCTGTCGGGGTAGCAGAGATCATCGRC 256
Dy	301 TGGGGGAGGAGCACTCTACAGGAGCAGCTGTCGGGGTAGCAGAGATCATCGRC 360
Qy	257 ACCCACAGTCTACACGCCAGATGGAGGGACATGCCCTGGAGCTGGAGGC 316
Dy	361 ACCCACAGTCTACATCATCCAGACTGGAGCGGATATGCCCTTGCTGGAGCTGGAGGC 420
Qy	317 CGGTGAAGGCTCCAGCCACGGTCAACAGGTACCCCTGGCTCAGAGACCTGC 376
Dy	421 CCCTGAACTCTTCAGCGCGTCACAGGTCATGCTCCCTGCTGGAGACCTGC 480
Qy	377 CCCGGGGATGCCGTGCGGGTCACTGGCTGGGGGAGTGGACATATGAGCCCTRC 436
Dy	481 CCCGGGGATGCCGTGCGGGTCACTGGCTGGGGGAGTGGACATATGAGCCCTRC 540
Qy	437 CACCGCCATTCCATTGAGCAGGTGAAGGTCCATAATGAAAACACATTGTGACG 496
Dy	541 CACCGCCATTCCCTGAGCAGGTGAAGGTCCCATATGAAAACACATTGTGACG 600
Qy	497 CAATAAACACCTTGCCCTACAGGGAGACCGACGGTCCATCGTGTGAGACATGC 556
Dy	601 CAATAAACACCTTGCCCTACAGGGAGACCGACGGTCCATCGTGTGAGACATGC 660
Qy	557 TGTGTGCGGGAAACCCGGAGGGACTATGCCAGGGGACTCGGAGGGCCCTGGT 616
Dy	661 TGTGTGCGGGAAACCCGGAGGGACTATGCCAGGGGACTCGGAGGGCCCTGGT 720
Qy	617 GCAAGGTGAATGGCACCTGGCTGTCAGCTGGG - CGAGGCTGTC 674
Dy	721 GCAGGTGATGGCACCCTGGCTACAGGGGGCTGGTCAGCTGGGAGGGGCTGTC 780
Qy	675 CGACGCCAACGGCTGGCATCTACACCGGTGCACTACTTGGACTGGATCCACCA 734
Dy	781 CGACGCCAACGGCTGGCATCTACACCGGTGCACTACTTGGACTGGATCCACCA 840
Qy	735 CTATGTCGCCAAAAGGGTGAAGGGCGCC 766
Dy	841 CTATGTCGCCAAAAGGGTGAATGCC 872
BASE COUNT	149
ORIGIN	a
Query Match	75.3%; Score 580.2; DB 14; length 904; Best Local Similarity 90.1%; Pred. No. 1.3e-112; Matches 666; Conservative 0; Mismatches 43; Indels 30; Gaps 3;
Qy	17 GAACTGCTGGGGTCAAGGGCCCGAGGAGCTGGCCCTGGCAGGTGAGAG 76
Dy	98 GCATGTTGGGGTCAGGGGCCAGGAGCTGGCCCTGGCAGGTGAGAG 157
Qy	77 TCCACGGCCCATACTGGACACTCTGGGGGCTCCATCCACCCCCAGTGGTGC 136
Dy	158 TCCGGGACGATCTGGTGCACCTCTGGGGCTCCATCCACCCCCAGTGGTGC 217
Qy	137 TGGGGGAGGAGCCTCACTACAGGACCTGTCGGCCGCTCAGGGTGAAC 196
Dy	218 TGGGGGAGGAGCCTCACTACAGGACCTGTCGGCCGCTCAGGGTGAAC 250
Qy	197 TGGGGGAGGAGCCTCACTACAGGACCTGTCGGCCGCTCAGGGTGAAC 256
Dy	251 TGGGGGAGGAGCCTCACTACAGGACCTGTCGGCCGCTCAGGGTGAAC 310
Qy	257 ACCCACAGTCTACACGCCAGATGGAGCCATGGGGAGGAGC 316

with tobacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA linker including EcoR I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained DNA vectors were used for transformation of competent cells E. coli Top10F<sup>+</sup> by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library.

**RESULT 3**

B0082929

LOCUS R0082929 502 bp mRNA linear EST 04-APR-2002

DEFINITION K-EST044841 S14K402 Homo sapiens cDNA clone S14K402-59-B11 5'

ACCESSION R0082929

VERSION R0082929.1 GI:19941018

KEYWORDS EST,

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1. (bases 1 to 502) Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R., Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and Kim, Y.S.

TITLE 21C Frontier Korean EST Project 2001

JOURNAL Unpublished (2002)

COMMENT Contact: Kim YS

Korea Research Institute of Bioscience & Biotechnology  
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel.: +82-42-860-4470  
Fax: +82-42-860-4409

Email: yonsung@mail.kribb.re.kr

Plate: 59 row: B column: 11  
High quality sequence stop: 502.

Location/Qualifiers 1. . 502

FEATURES source

/organism="Homo sapiens"  
/clone="S14K402-59-B11"  
/clone\_1ib="S14K402"  
/cell\_line="R402"  
/lab\_host="Top10F"  
/notes="Organ: Stomach; Vector: pTZ18RPI; Site\_1: EcoRI; Site\_2: NotI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped"

**RESULT 4**

EG542089

LOCUS B0542089 736 bp mRNA linear EST 03-APR-2001

DEFINITION 602571309f1 NIH\_MGC\_77 Homo sapiens cDNA clone IMAGE:4695458 5'

ACCESSION B0542089

VERSION B0542089.1 GI:13534322

KEYWORDS EST,

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1. (bases 1 to 736) NIH-MGC http://mgc.nci.nih.gov/

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT									
<p>Contact: Robert Strausberg, Ph.D.  Email: cagpb-r@mail.nih.gov  Tissue Procurement: CLONTECH Laboratories, Inc.  cDNA Library Preparation: CLONTECH Laboratories, Inc.  DNA Sequencing by: Incyte Genomics, Inc.  Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  <a href="http://Image.llnl.gov">http://Image.llnl.gov</a>  Plate: LUCM1521 row: b column: 03  High quality sequence stop: 725.</p>									
FEATURES	source	Location/Qualifiers							
		1. . 736							
BASE COUNT		/organism="Homo sapiens" /lab_host="DH10B ("T1 phage-resistant")" /note="Organ: lung; Vector: pBR-L1B (Clontech); Site_1: SfiI (ggccgcctggcc); Site_2: SfiI (ggccatttggcc); 5'; and 3' adaptors were used in cloning, as follows: 5' adaptor sequence: 5'-CACGGCCATATTGCC-3', and 3' adaptor sequence: 5'-ATTCAGGCGAGGCGGCCACATG-dT(30)BN-3', (where B = A, C, or G and N = C, G, or T). Average insert size 1.9 kb (range 0.5-4.0 kb). 12/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC library."							
ORIGIN		138 a 236 c 242 g 120 t							
Query Match		60.8%; Score 468.4; DB 12; Length 736;							
Best Local Similarity		90.9%; Pred. No. 5.2e-19; Mismatches 46; Indels 11; Gaps 6;							
Matches 567; Conservative 0; Mismatches 46;		/db_xref="taxon:9606" /clone="IMAGE:469508" /clone_lib="NIH_MGC_77"							
Qy	17	GAATCTCGGGGTCTAGGAGGCCAACGGCCAAGGGCCCTGGAGGTAGCCCTGAGAG							
Db	113	GCATCGTTGGGGTCAGGAGGCCAGGAACAGTGCCCTGGAGGTAGCCCTGAGAG							
Qy	77	TCCACGGCCATACTGGATGACTCTGGGGGGCTCCCATCACCCCCAGGGTG							
Db	173	TCCACGCCCATACTGGATGACTCTGGGGGGCTCCCATCACCCCCAGGGTG							
Qy	137	TGACCGCCGCGCGCTGGGGACCGAGGCAAGAAGTCGGCCGCTCAGGGTG							
Db	233	TGACCGCAGCGCACTGGGGACCGAGGCAAGGTCAAAGGATCTGGCCGCTCAGGGTG							
Qy	197	TGGGGAGACGACCCACTCTAACCGAGGACAGCTGCTGGGTCAAGGATCATGTC							
Db	293	TGCGGGAGACGACCCACTCTAACCGAGGACATCGCCCTGCGTGGAGGAGC							
Qy	257	ACCCAGCTCTACAGGCCAGATGGAGGGACATCGCCCTGCGTGGAGGAGC							
Db	353	ACCCACGTTCTACAGGCCAGATGGAGGGACATCGCCCTGCGTGGAGGAGC							
Qy	317	CGGTGAGGGTCTCCAGGCCACCTGCAACGGTCACCCCTGCCCCCTGCTAGAGCTTC							
Db	413	CGGTGAGGGTCTCCAGGCCACCTGCAACGGTCACCCCTGCTAGAGCTTC							
Qy	377	CCCC-GGGGATGCCGCGCTGGTCACTGGGCTGGGGGATGGACATGGAGGCCCTC							
Db	473	CCCCGGGATGCCGCGCTGGTCACTGGGCTGGGGGATGGACATGGAGGCCCTC							
Qy	436	CCACCGCCATTCTCTGAGGAG-TGAAGTCCCATAATGAAACACATGGTGA							
Db	533	CCACGGCATTCCTCTGAGCAGGTGAGGTCGCCATATGAAACACATGGTGA							
Qy	495	CGCAAATACCGCTGGGCCT-ACACGGGAGAC-ACGCGCCGCTCGCGTGCAGGAC							
Db	593	CGCAAATAACGACTTGGGCCTAACGGGAGACAGGTCGCCATCTGGGTAAC							
Qy	553	ATGCT----TGTGGGGAGACACCGGGAGGACTCATGCGAGGCTGGTGAAGGTCCCATAATGG							
BASE COUNT		178 a 335 c 240 g 164 t							
FEATURES	source	Location/Qualifiers							
		1. . 917							
Query Match		57.6%; Score 444.4; DB 14; Length 917;							
Best Local Similarity		98.5%; Pred. No. 6.4e-84; Mismatches 0; Indels 1; Gaps 1;							
Matches 459; Conservative 0; Mismatches 0; Indels 1; Gaps 1;		/db_xref="taxon:619508" /clone="IMAGE:619508 5'", mRNA sequence.							
RESULT 5	Qy	606	GCCCCCTGTCACAAGGTGAATTGG 629	LOCUS	BQ721078	917 bp	mRNA	1 linear	EST 16-JUL-2002
DEFINITION			AGENCOURT-8109065 Lupsk1_sympathetic_trunk	ORGANISM	Homo sapiens				EST cdNA clone
ACCESSION			BQ721078	KEYWORDS					
VERSION			BQ721078.1	COMMENT					
SOURCE				REFERENCE					
ORGANISM				AUTHORS					
Homo sapiens				TITLE					
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo				JOURNAL					
(bases 1 to 917)				COMMENT					
1 (bases 1 to 917)				CONTACT					
NIH_MGC http://mgc.ncbi.nlm.nih.gov/.				Email:	cagpb-r@mail.nih.gov				
National Institutes of Health, Mammalian Gene Collection (MGC)				Tissue Procurement:	Dr. James R. Lupski				
Unpublished (1999)				cDNA Library Preparation:	Life Technologies, Inc.				
Contact: Robert Strausberg, Ph.D.				cDNA Library Arranged by:	The I.M.A.G.E. Consortium (LLNL)				
				DNA Sequencing by:	Agencourt Bioscience Corporation				
				Clone distribution:	MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:				
RESULT 5	Qy	606	GCCCCCTGTCACAAGGTGAATTGG 629	LOCUS	BQ721078	917 bp	mRNA	1 linear	EST 16-JUL-2002
DEFINITION			AGENCOURT-8109065 Lupsk1_sympathetic_trunk	ORGANISM	Homo sapiens				EST cdNA clone
ACCESSION			BQ721078	KEYWORDS					
VERSION			BQ721078	COMMENT					
SOURCE				REFERENCE					
ORGANISM				AUTHORS					
Homo sapiens				TITLE					
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo				JOURNAL					
(bases 1 to 917)				COMMENT					
1 (bases 1 to 917)				CONTACT					
NIH_MGC http://mgc.ncbi.nlm.nih.gov/.				Email:	cagpb-r@mail.nih.gov				
National Institutes of Health, Mammalian Gene Collection (MGC)				Tissue Procurement:	Dr. James R. Lupski				
Unpublished (1999)				cDNA Library Preparation:	Life Technologies, Inc.				
Contact: Robert Strausberg, Ph.D.				cDNA Library Arranged by:	The I.M.A.G.E. Consortium (LLNL)				
				DNA Sequencing by:	Agencourt Bioscience Corporation				
				Clone distribution:	MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:				
RESULT 5	Qy	606	GCCCCCTGTCACAAGGTGAATTGG 629	LOCUS	BQ721078	917 bp	mRNA	1 linear	EST 16-JUL-2002
DEFINITION			AGENCOURT-8109065 Lupsk1_sympathetic_trunk	ORGANISM	Homo sapiens				EST cdNA clone
ACCESSION			BQ721078	KEYWORDS					
VERSION			BQ721078	COMMENT					
SOURCE				REFERENCE					
ORGANISM				AUTHORS					
Homo sapiens				TITLE					
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo				JOURNAL					
(bases 1 to 917)				COMMENT					
1 (bases 1 to 917)				CONTACT					
NIH_MGC http://mgc.ncbi.nlm.nih.gov/.				Email:	cagpb-r@mail.nih.gov				
National Institutes of Health, Mammalian Gene Collection (MGC)				Tissue Procurement:	Dr. James R. Lupski				
Unpublished (1999)				cDNA Library Preparation:	Life Technologies, Inc.				
Contact: Robert Strausberg, Ph.D.				cDNA Library Arranged by:	The I.M.A.G.E. Consortium (LLNL)				
				DNA Sequencing by:	Agencourt Bioscience Corporation				
				Clone distribution:	MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:				
RESULT 5	Qy	606	GCCCCCTGTCACAAGGTGAATTGG 629	LOCUS	BQ721078	917 bp	mRNA	1 linear	EST 16-JUL-2002
DEFINITION			AGENCOURT-8109065 Lupsk1_sympathetic_trunk	ORGANISM	Homo sapiens				EST cdNA clone
ACCESSION			BQ721078	KEYWORDS					
VERSION			BQ721078	COMMENT					
SOURCE				REFERENCE					
ORGANISM				AUTHORS					
Homo sapiens				TITLE					
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo				JOURNAL					
(bases 1 to 917)				COMMENT					
1 (bases 1 to 917)				CONTACT					
NIH_MGC http://mgc.ncbi.nlm.nih.gov/.				Email:	cagpb-r@mail.nih.gov				
National Institutes of Health, Mammalian Gene Collection (MGC)				Tissue Procurement:	Dr. James R. Lupski				
Unpublished (1999)				cDNA Library Preparation:	Life Technologies, Inc.				
Contact: Robert Strausberg, Ph.D.				cDNA Library Arranged by:	The I.M.A.G.E. Consortium (LLNL)				
				DNA Sequencing by:	Agencourt Bioscience Corporation				
				Clone distribution:	MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:				
RESULT 5	Qy	606	GCCCCCTGTCACAAGGTGAATTGG 629	LOCUS	BQ721078	917 bp	mRNA	1 linear	EST 16-JUL-2002
DEFINITION			AGENCOURT-8109065 Lupsk1_sympathetic_trunk	ORGANISM	Homo sapiens				EST cdNA clone
ACCESSION			BQ721078	KEYWORDS					
VERSION			BQ721078	COMMENT					
SOURCE				REFERENCE					
ORGANISM				AUTHORS					
Homo sapiens				TITLE					
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo				JOURNAL					
(bases 1 to 917)				COMMENT					
1 (bases 1 to 917)				CONTACT					
NIH_MGC http://mgc.ncbi.nlm.nih.gov/.				Email:	cagpb-r@mail.nih.gov				
National Institutes of Health, Mammalian Gene Collection (MGC)				Tissue Procurement:	Dr. James R. Lupski				
Unpublished (1999)				cDNA Library Preparation:	Life Technologies, Inc.				
Contact: Robert Strausberg, Ph.D.				cDNA Library Arranged by:	The I.M.A.G.E. Consortium (LLNL)				
				DNA Sequencing by:	Agencourt Bioscience Corporation				
				Clone distribution:	MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:				
RESULT 5	Qy	606	GCCCCCTGTCACAAGGTGAATTGG 629	LOCUS	BQ721078	917 bp	mRNA	1 linear	EST 16-JUL-2002
DEFINITION			AGENCOURT-8109065 Lupsk1_sympathetic_trunk	ORGANISM	Homo sapiens				EST cdNA clone
ACCESSION			BQ721078	KEYWORDS					
VERSION			BQ721078	COMMENT					
SOURCE				REFERENCE					
ORGANISM				AUTHORS					
Homo sapiens				TITLE					
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo				JOURNAL					
(bases 1 to 917)				COMMENT					
1 (bases 1 to 917)				CONTACT					
NIH_MGC http://mgc.ncbi.nlm.nih.gov/.				Email:	cagpb-r@mail.nih.gov				
National Institutes of Health, Mammalian Gene Collection (MGC)				Tissue Procurement:	Dr. James R. Lupski				
Unpublished (1999)				cDNA Library Preparation:	Life Technologies, Inc.				
Contact: Robert Strausberg, Ph.D.				cDNA Library Arranged by:	The I.M.A.G.E. Consortium (LLNL)				
				DNA Sequencing by:	Agencourt Bioscience Corporation				
				Clone distribution:	MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:				
RESULT 5	Qy	606	GCCCCCTGTCACAAGGTGAATTGG 629	LOCUS	BQ721078	917 bp	mRNA	1 linear	EST 16-JUL-2002
DEFINITION			AGENCOURT-8109065 Lupsk1_sympathetic_trunk	ORGANISM	Homo sapiens				EST cdNA clone
ACCESSION			BQ721078	KEYWORDS					
VERSION			BQ721078	COMMENT					
SOURCE				REFERENCE					
ORGANISM				AUTHORS					
Homo sapiens				TITLE					
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo				JOURNAL					
(bases 1 to 917)				COMMENT					
1 (bases 1 to 917)				CONTACT					
NIH_MGC http://mgc.ncbi.nlm.nih.gov/.				Email:	cagpb-r@mail.nih.gov				
National Institutes of Health, Mammalian Gene Collection (MGC)				Tissue Procurement:	Dr. James R. Lupski				
Unpublished (1999)				cDNA Library Preparation:	Life Technologies, Inc.				
Contact: Robert Strausberg, Ph.D.				cDNA Library Arranged by:	The I.M.A.G.E. Consortium (LLNL)				
				DNA Sequencing by:	Agencourt Bioscience Corporation				
				Clone distribution:	MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:				
RESULT 5	Qy	606	GCCCCCTGTCACAAGGTGAATTGG 629	LOCUS	BQ721078	917 bp	mRNA	1 linear	EST 16-JUL-2002
DEFINITION			AGENCOURT-8109065 Lupsk1_sympathetic_trunk	ORGANISM	Homo sapiens				EST cdNA clone
ACCESSION			BQ721078	KEYWORDS					
VERSION			BQ721078	COMMENT					</td

QY 479 AAACACATTGTGAGCCAAATTCACACTTGCGCTTACACGGAGACGAGCTCCGA 538  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 181 AAACACATTGTGAGCCAAATTCACACTTGCGCTTACACGGAGACGAGCTCCGA 240  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 QY 539 TCGTCGTGAGCA-CATGCTGTGAGCCGGAAACACCGGGACATCAGCAGGAGC 597  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 241 TCGTCGTGAGCA-CATGCTGTGAGCCGGAAACACCGGGACATCAGCAGGAGC 300  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 QY 598 TCCGAGGGCCCTGTTGCAAGGGAATGGACCTGCTGAGGGCGGCTGAGGAGC 657  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 301 TCCGAGGGCCCTGTTGCAAGGGAATGGACCTGCTGAGGGCGGCTGAGGAGC 360  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 QY 658 TGGGGGAGGGCTGTCAGGCCAACGGCTTGCACTACACCGGTTGACCTACTAC 717  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 361 TGGGGGAGGGCTGTCAGGCCAACGGCTTGCACTACACCGGCTGAGGAGC 420  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 QY 718 TTGACTGATCCACACTATGTCCTCCAAAAGGGTGAAGCGGC 763  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 421 TGGGACTGGTCCACACTATGTCCTCCAAAAGGGTGAAGCGGC 466  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 RESULT 6  
 BI13250 BI13250 1000 bp mRNA linear EST: 14-AUG-2001  
 DEFINITION 60298390F1 NCL\_CGAP\_Lu33 MUS musculus cDNA clone IMAGE:5142478 5',  
 mRNA sequence.  
 ACCESSION BI13250  
 AUTHORS NIH-MGC http://mgc.ncbi.nih.gov/  
 VERSION BI13250.1 GI:15174173  
 KEYWORDS EST,  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 1000)  
 NIH-MGC http://mgc.ncbi.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbsr@mail.nih.gov  
 Tissue Procurement: Gilbert Smith, Ph.D.  
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: NCI-CGAP clone distribution information can be found through the T.M.A.G.E. Consortium/LINN at:  
 http://image.llnl.gov  
 Plate: LIAM11351 row: c column: 23  
 High quality sequence start: 32  
 High quality sequence stop: 849.  
 Location/Qualifiers  
 FEATURES source  
 1. .1000  
 /organism="Mus musculus"  
 /strain="CZCH 11"  
 /clone="IMAGE:5142478"  
 /clone\_1b="NCI\_CGAP\_Lu33"  
 /tissue\_type="pooled lung tumors"  
 /lab\_host="DIBOB (phage-resistant)"  
 /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site\_1: NotI; Site\_2: EcoRI; 1st strand cDNA was prepared from mRNA obtained from pooled lung tumors with a Not I - oligo(dT) primer [5, TCTAACCATCTGAGTGGACGGGCCCTGTTTTTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo. "  
 BASE COUNT 202 a 294 c 274 g 229 t 1 others  
 ORIGIN Query Match

Best Local Similarity 77.2%; Pred. No. 3.2e-79; Matches 564; Conservative 0; Mismatches 163; Indels 4; Gaps 4;

QY 17 GAATGTCGGGGTCAGGAGGCCAGGAGCAAGTGGCCCTGGCGAGCTGAGGAG 76  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 73 GCATCGTGGAGGACATGAGCTTGAGTAGTGGCCCTGCGAGTGAAGCTGAGAT 132  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 QY 77 TCCACGGCCATACCTGAGGACACTGCTGAGGAGCTGAGTAGTGGCCCTGCGAGTGAAGCTGAGAT 136  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 133 TTAATTAACATCTGGATACATTCTGGGAGGCTCTCCACCCAGTGGTGC 192  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 QY 137 TGACGCCSGGCTTGCGTGGAGCTGAGGATGTGGCGCCCTGAGGTCAC 196  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 193 TCACTGCGCACACTGAGGAGCCACATCAAAGGCCAACAGCTCTCCGGGTGAGC 252  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 QY 197 TGGGGAGCAGCACCTACTACAGGACCGCTGCTGGCAGCAGGATATCGTC 256  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 253 TCCGGAGGAGCTATGAGGACGACGCTGAGGAGCTGGGAGCTGGGAGTC 312  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 257 ACCCACAGTGTACACGCCAGATCGAGGAGCATGCCCTGCTGGAGCTGGAGGAC 316  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 313 ACCCCACTATACACGGCCAGGGTGGCCAGCTGGAGCTGGGAGCTGGTC 372  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 QY 317 CGGTGAAGGTTCCAGCCACTGTCACCGTACCCCTGCCCTGCCCTCAGAACCTTC 376  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 373 CTGTTGAACTTCTCCACCCATATCCACCCATATCCCTGCCCTGCCCTGAGACTTCC 432  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 QY 377 CCCCGGGATGGCGTGGCTGGTCACTGGCTGGGATGAGCATGAGGCCCTCC 436  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 433 CCCCTGGGACATCTGGGAGCTGGTGGAGCTGGGAGCTGGGAGCTGG 492  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 QY 437 CACGCCATTCTCTGAAGCAGGAGTCCCATATGGAAAACCACATTGTA-C 495  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 493 CACCTCTTATCTCTGAGACCAAGTGAAGGTCCCATGTTGGAARACAGCCGTG 552  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 QY 496 GCAAAATACACCTTGCCCTAACGGAGAACGGAGCTGGTGGAGCTGG 555  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 553 GGAAGTAACCAACACTGGCCCTACACGGAGATGATTTCGATGAGTC 612  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 QY 556 CTGGTGGGGACACCGGAGGACTATGCCAGGGACTCGCAGGGGACTCGGGGGCTGG 615  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 613 CTGTTGCTGGAAATACCGGAGACTCTGGGAGCTGGGGCTGG 672  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 QY 616 TGCAGGTGAATGGACCTGGCTGAGGGGGCTGGTA-GCCTGGGGAGGGCTGG 674  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 673 TGCAGATGAGGTACCTGGCTGGAGGAGTGGTCACTGGGGGGTGG 732  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 QY 675 CAGGCC-ACCGCCCTGACATCACACCGGTGTCACC-TACTACTGGACTGATCCAC 732  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 733 ACAGCCCAACAAAGCTGGCATCTACACCCGGTGGACATACTACTAGCTGGATCCAC 792  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 QY 733 CACTATGCC 743  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 793 GCTTATGTCGCC 803  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 7  
 AA131142  
 LOCUS AA131142  
 DEFINITION AA131142.r1 Soares-pregnant uterus\_nbpu Homo sapiens cDNA clone IMAGE:503497 5', similar to gb:M30038 ALPHA-TRYPTASE PRECURSOR (HUMAN), mRNA sequence.  
 ACCESSION AA131142  
 VERSION AA131142.1 GI:1692796  
 KEYWORDS EST,  
 SOURCE  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 639)  
 AUTHORS Hillier, L., Lennon, G., Becker, M.F., Bonaldo, M.F., Chiapelli, B., Chissoe, S., Dietrich, N., Dubreuil, T., Favello, A., Gish, W., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore

COMMENT	Contact: Wilson RK Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 266 1800 Fax: 314 286 1810 Email: est@wustl.edu	Db 420 CTACACCCGTGTCACCTACTTGGAACTGGATCCACCANTATGTCACCAAAAGCCGT 479 BGC697540 811 bp mRNA linear EST 07-MAY-2001 LOCUS BG697540 811 bp mRNA linear EST 07-MAY-2001 DEFINITION 602660760F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4803879 5', JOURNAL BG697540 MEDLINE BG697540.1 GI:1396867 97044478
FEATURES	This clone is available royalty-free through LLNL; contact the IMAGE Consortium ( <a href="http://image.llnl.gov">http://image.llnl.gov</a> ) for further information. Insert Length: 894 Std Error: 0.00 Seq primer: -28M13 rev2 from Amersham High quality sequence stop: 349. Location/Qualifiers	
source	1. .639 /organism="Homo sapiens" /ab_xref="GB:3808485" /db_xref="taxon:9606" (clone="IMAGE:503497" /clone.lib=Soares_pregnant_uterus_NDHPHU" /sex="female" /dev_stage="adult" /lab_host="DH10B" /note="Organ: uterus; Vector: pT7T3-Pac; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', double-stranded cDNA was ligated to T7PPT 3']" (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library constructed by M. Fatima Bonaldo." went through one round of normalization. Library constructed by M. Fatima Bonaldo."	
BASE COUNT	120 a 206 c 175 g 120 t 18 others ORIGIN	
Query Match	54.5%; Score 420.4; DB 9; Length 639;	
Best Local Similarity	94.0%; Pred. No. 6.9e-79;	
Matches	450; Conservative 0; Mismatches 27; Indels 2; Gaps 2;	
Db	1 CCAGACTGGAGCGGATATGCCCTTGCTGGAGCTGGAGGCCGATCTCCAGCC 305	
Qy	376 CGTCCACACGGTACCCCTGCCCTCAGAGAACCTCCCCGGGATGCCGTG 395	
Db	61 CGTCACACGGTCATGCTGCCCTGCCCTGGAGAACCTTCCCGGGATGCCGTG 120	
Qy	396 GGTCCTCTGGTGGCGCATGTGGACATGATGAGGCCCTCCACGCCATTCTCTGAA 455	
Db	121 GGTCACCTGGTGGGCCGATGTGGACAAATGATGAGGCCCTCCACGCCATTCCCTGA 180	
Qy	456 GCACGTGAAGGTCCCATAATGGAAACCACTTGTGGCAAATAACCACTTGGCC 515	
Db	181 GCAGGTGAAGTCCCATATGGAAACCACTTGTGGCAAATAACCACTTGGCC 240	
Qy	516 CTACCGGGGACGAGCTGGCAGTGTGGACACATGCTGTGGGGACACCCG 575	
Db	241 CTACCGGGGACGAGCTGGCAGTGTGGACACCCG 300	
Qy	576 GAGGACTCATGCCGGGCACTCGGGAGGCCCTGGTGCAGGTGAAATGGCACCTG 635	
Db	301 GAGAGCTCATGCCGGGCACTCGGGAGGCCCTGGTGCAGGTGAAATGGCACCTG 360	
Qy	636 GCTGGAGGGGGGGGGGGTCACTGGGGGGGGTGTGGCCASGCCAACGGGCTGGCAT 695	
Db	361 GCTACAGGGGGGGGGTCACTGGGAGGGGGTGTG-CAGTCCTAACCGGCCCTGGCAT 419	
Qy	696 CTACACCCGGTCACTACTTGTG- ACTGGATCACCACATATGCCCAAAGCCGT 754	
FEATURES	ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Primates; Catarrhini; Hominida; Homo. REFERENCE 1 (bases 1 to 811) AUTHORS NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> TITLE National Institutes of Health, Mammalian Gene Collection (MGC) JOURNAL Unpublished (1999) COMMENT Contact: Robert Strausberg, Ph.D. Email: cgbps-r@mail.nih.gov Tissue Procurement: James Cleaver, M.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The T.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the T.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LLM1070 Row: g column: 16 High quality sequence stop: 781. Location/Qualifiers	
BASE COUNT	1. .811 /organism="Homo sapiens" /ab_xref="taxon:9606" (clone="IMAGE:4803879" /clone.lib=NCI_CGAP_Skn3" /lab_host="DH10B (T1 phage-resistant)" /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: Not I; Site_2: Sall; Cloned unidirectionally. Primer: Oligo dT. Average insert size: 1.5kb. Library constructed by Life Technologies. Note: this is a NCI_CGAP Library." ORIGIN	
Query Match	52.1%; Score 401.4; DB 12; Length 811;	
Best Local Similarity	98.5%; Pred. No. 7.6e-75;	
Matches	405; Conservative 0; Mismatches 6; Indels 0; Gaps 0;	
Db	17 GAATCGTCGGGGTCAGGGGCCAGAGCAATGGGGCTGGCAGGTGGAG 76	
Qy	98 GCACTGGTTGGGGTCAGGGGCCAGAGCAATGGGGCTGGCAGGTGGAG 157	
Db	77 TCCACGGCCATACTGGATGCCACTCTGGGGGAGCTCCCTCACCCCTACTGGGC 136	
Qy	158 TCCACGGCCCATACTGGATGCCACTCTGGGGGCTCCCTCACCCCTACTGGGC 217	
Db	137 TACGGCCGGGCGGCCAGGAGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGG 196	
Qy	218 TGACCGCACGCCACTGGGGAGGCCAGGAGCTGGGGCTGGGGCTGGGGCTGGGG 277	
Db	197 TGCGGGAGCACCTCTACCTACGGAGGCCAGGAGCTGGGGCTGGGGCTGGGG 256	
Qy	278 TCGGGAGCACCTCTACCTACGGAGGCCAGGAGCTGGGGCTGGGGCTGGGGCTGGGG 337	
Db	257 ACCCACAGTCTACCGGCCAGTGGAGGGACATGCCCTCTGGAGCTGGAGGC 316	
Qy	338 ACCCACAGTCTACCGGCCAGTGGAGGGACATGCCCTCTGGAGCTGGAGGC 397	
Db	317 CGGTGAAGTCTCCAGGCCACACGGTACCTGGCCCTGGCTAGAGACCTTC 376	

Db	398	CGGTGAAGTCCTCCAGCCACGGTCAACACGGTCACCGTCACCTGCTGCCCTGCAGAGACCTCC	457
Qy	377	CCCGGGATGCGCTGCTGGTCACTGGCTGGGGGAGATGGACAATGATG	427
Db	458	CCCCGGGATGCCGTGCTGGTCACTGGCTGGGGGATGGACATGATG	508
RESULT 9			
LOCUS	BG483591	678 bp	mRNA
DEFINITION	602503109FL NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4616715 5'	linear	EST 21-MAR-2001
VERSION	BG483591	1	EST.
KEYWORDS	mRNA sequence.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. (bases 1 to 678)		
AUTHORS	NH_MGC http://mgc.ncbi.nih.gov/		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabps-r@mail.nih.gov		
FEATURES	source		
	1.	678	
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:4616715"		
	/clone.lib="NIH_MGC_77"		
	/lab_host="DHL0B (T1 phage-resistant)"		
	/note="Organ: lung; Vector: PDNR-LIB (Clontech); Site_1: SFI1 (ggccgcgtcgcc); Site_2: SFI1 (ggccatggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGCGC-3', and 3' adaptor sequence: 5'-ATTCAGASGCCGAGCGGCCGAGATG-dT(30)BN-3', (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.9 kb (range 0.5-4.0 kb). 12/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."		
BASE COUNT	a	129	
ORIGIN	c	224	
	g	215	
	t	110	
FEATURES	source		
	1.	.512	
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone.lib="NIH_MGC_77"		
	/clone_no="bx01h02"		
	/BX,"		
	/tissue_type="Iris"		
	/dev_stage="adult"		
	/lab_host="EMDH10"		
	/note="Organ: Eye; Vector: pCMVSPORT6; Post-mortem iris tissue was pooled from 10 individuals ranging in age from 4-80 years and RNA was extracted. From this pooled sample an aliquot of 60ug of total RNA yielded 2.17ug of mRNA. A directionally cloned cDNA library was constructed using the pCMVSPORT6 vector was constructed at Life Technologies, essentially following the protocols of the SuperScript Plasmid System full details of which are contained in the manufacturer's Instruction manual (http://www.lifetech.com/). First strand synthesis was carried out using a Not I primer-adapter [5'-pGACTGTTCTACGGAGCAGCTACAGGAGATCATCG		
RESULT 10			
LOCUS	BF724180	512 bp	mRNA
DEFINITION	bx01h02.Y2 Human Iris cDNA (Un-normalized, unamplified): BX_Homo_sapiens cDNA clone bx01h02 5', mRNA sequence.	linear	EST 05-JAN-2001
VERSION	BF724180		
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. (bases 1 to 512)		
AUTHORS	Wistow,G.J., Bernstein,S., Behar,A. and Smith,D.		
TITLE	NEIBANK: EST analysis and bioinformatics for ocular genomics		
JOURNAL	Invest. Ophthalmol. Vis. Sci. 41, (2000) In press		
COMMENT	Contact: Wistow G Section on Molecular Structure and Function National Eye Institute Tel: 301 496 0078 Fax: 301 492 3452 Email: graeme@helix.nih.gov Plate: 01 column: 02 row: h		
FEATURES	source		
	Seq primer: M13RPI reverse primer (ABI).		
	1.	.512	
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone.lib="Human Iris cDNA (Un-normalized, unamplified): BX,"		
	/tissue_type="Iris"		
	/dev_stage="adult"		
	/lab_host="EMDH10"		
	/note="Organ: Eye; Vector: pCMVSPORT6; Post-mortem iris tissue was pooled from 10 individuals ranging in age from 4-80 years and RNA was extracted. From this pooled sample an aliquot of 60ug of total RNA yielded 2.17ug of mRNA. A directionally cloned cDNA library was constructed using the pCMVSPORT6 vector was constructed at Life Technologies, essentially following the protocols of the SuperScript Plasmid System full details of which are contained in the manufacturer's Instruction manual (http://www.lifetech.com/). First strand synthesis was carried out using a Not I primer-adapter [5'-pGACTGTTCTACGGAGCAGCTACAGGAGATCATCG		
RESULT 11			
LOCUS	BX01H02	512 bp	mRNA
DEFINITION	bx01h02.Y2 Human Iris cDNA (Un-normalized, unamplified): BX_Homo_sapiens cDNA clone bx01h02 5', mRNA sequence.	linear	EST 05-JAN-2001
VERSION	BX01H02		
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. (bases 1 to 512)		
AUTHORS	Wistow,G.J., Bernstein,S., Behar,A. and Smith,D.		
TITLE	NEIBANK: EST analysis and bioinformatics for ocular genomics		
JOURNAL	Invest. Ophthalmol. Vis. Sci. 41, (2000) In press		
COMMENT	Contact: Wistow G Section on Molecular Structure and Function National Eye Institute Tel: 301 496 0078 Fax: 301 492 3452 Email: graeme@helix.nih.gov Plate: 01 column: 02 row: h		
FEATURES	source		
	Seq primer: M13RPI reverse primer (ABI).		
	1.	.512	
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone.lib="Human Iris cDNA (Un-normalized, unamplified): BX,"		
	/tissue_type="Iris"		
	/dev_stage="adult"		
	/lab_host="EMDH10"		
	/note="Organ: Eye; Vector: pCMVSPORT6; Post-mortem iris tissue was pooled from 10 individuals ranging in age from 4-80 years and RNA was extracted. From this pooled sample an aliquot of 60ug of total RNA yielded 2.17ug of mRNA. A directionally cloned cDNA library was constructed using the pCMVSPORT6 vector was constructed at Life Technologies, essentially following the protocols of the SuperScript Plasmid System full details of which are contained in the manufacturer's Instruction manual (http://www.lifetech.com/). First strand synthesis was carried out using a Not I primer-adapter [5'-pGACTGTTCTACGGAGCAGCTACAGGAGATCATCG		



SIMPSON, A.J.	TITLE	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL	PROC. NATL. ACAD. SCI. U.S.A. 97 (7), 3491-3496 (2000)	
MEDLINE	20202663	
COMMENT	Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antônio Prudente 109, 4 andar, 01509-010, São Paulo-SP, Brazil Tel: +55-11-2704922 Fax: +55-11-2707001	
	Email: asimpson@ludwig.org.br	
	This sequence was derived from the FAPESP/LICR Human Cancer Genome project. This entry can be seen in the following URL ( <a href="http://www.ludwig.org.br/scripts/gethtml2.p1?t1=CM3&amp;t2=EN0077-181100-489-e06&amp;t3=2000-11-18&amp;t4=1">http://www.ludwig.org.br/scripts/gethtml2.p1?t1=CM3&amp;t2=EN0077-181100-489-e06&amp;t3=2000-11-18&amp;t4=1</a> )	
FEATURES	Seq. Primer: puc 18 forward High quality sequence start: 28 High quality sequence stop: 376. Location/Qualifiers	
source	1. .376 <organism="Homo sapiens"> <db_xref="taxon:9606"> <clone_1lib="EN0077"> <dev_stage="Adult">	
BASE COUNT	Best Local Similarity 96.5%; Pred. No. 4.1e-64; Matches 358; Conservative 0; Mismatches 13; Indels 0; Gaps 0;	
ORIGIN	Query Match 45.4%; Score 350.2; DB 12; Length 376; Site_1: Small; Site_2: SmtI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."	
	62 a 115 c 121 g 78 t	
FEATURES	Location/Qualifiers	
source	1. .692 <organism="Homo sapiens"> <db_xref="taxon:9606"> <clone_1lib="IMAGE:5862333"> <clone_1lib="NCI_CGAP_D10"> <tissue_type="Lung_Focal_Fibrosis"> <dev_stage="Adult"> <lab_host="DH10B (Life Technologies)"> <note="organ: Lung; vector: pT73-Pac (Pharmacia) with a modified polylinker; Site_1: Eco R I; Site_2: Not I;"> NCI_CGAP_D10 is a cDNA library containing the following tissue(s): A pool of lung Focal Fibrosis. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an Eco R I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first strand cDNA contains a library tag sequence that is located between the Not I site and the ATAGGGTC. (dP18 tail). The sequence tag for this library is TAG LIB-D10-H-D10 TAG TISSUE=lung with fibrosis TAG SEQ=ATACGGGC"	
BASE COUNT	130 a 182 c 242 g 137 t 1 others	
ORIGIN	Query Match 45.3%; Score 349.6; DB 14; Length 692; Site_1: Small; Site_2: SmtI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."	
	195 GGCTCTCCAGGCCAGTCCACAGGTCACTGGCCCTCTGCTCTAGAGACCTCCCCCGG 383 Db 255 GTTCGACACGCCAGATCGGAGCATCGCCCTGCTGGAGGAGCCGTCGA 196	
ORIGIN	Query Match 45.3%; Score 349.6; DB 14; Length 692; Site_1: Small; Site_2: SmtI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."	
	324 GGTCCTCCAGGCCAGTCCACAGGTCACTGGCCCTCTGCTCTAGAGACCTCCCCCGG 383 Db 195 GGCTCTCCAGGCCAGTCCACAGGTCACTGGCCCTCTGCTCTAGAGACCTCCCCCGG 136	
ORIGIN	Query Match 45.3%; Score 349.6; DB 14; Length 692; Site_1: Small; Site_2: SmtI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."	
	384 GATGCCGTCCTGGCTACTCGCTGGCGATGTGGACATGTAGAGCCTCCACCCCG 443 Db 135 GATGCCGTCCTGGCTACTCGCTGGCGATGTGGACATGTAGAGCCTCCACCCCG 76	
ORIGIN	Query Match 45.3%; Score 349.6; DB 14; Length 692; Site_1: Small; Site_2: SmtI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."	
	444 ATTCTCTCTGAAAGGAGGAGGCTCCATAATGGAAACACATTGTGACCCAAATA 503 Db 75 ATTCTCTCTGAAAGGAGGCTCCATAATGGAAACACATTGTGACCCAAATA 16	
ORIGIN	Query Match 45.3%; Score 349.6; DB 14; Length 692; Site_1: Small; Site_2: SmtI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."	
	504 CCACCTTGCG 514 Db 15 GCACATTAAGTC 5	
RESULT	13	
DEFINITION	BM989945/C	
LOCUS	BM989945	mRNA linear EST 17-JUN-2002
IMAGE	:5862333 3' mRNA sequence.	
ACCESSION	BA989945	
VERSION	Bm989945.1	GI:19709334
KEYWORDS		EST.
SOURCE		human.
ORGANISM	Homo sapiens	
JOURNAL	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.	
COMMENT		1 (bases 1 to 692)
REFERENCE	NCI-CGAP <a href="http://www.ncbi.nlm.nih.gov/ncicgap">http://www.ncbi.nlm.nih.gov/ncicgap</a> .	
AUTHORS	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index.	
TITLE	Unpublished (1997)	
COMMENT	Contact: Robert Strausberg, Ph.D.	
REFERENCE	Email: cggps@mail.nih.gov	
AUTHORS	Tissue Procurement: Dr. Jose Mercuende CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: Clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a>	
POLYA=Yes		
FEATURES	Location/Qualifiers	
source	1. .692 <organism="Homo sapiens"> <db_xref="taxon:9606"> <clone_1lib="IMAGE:5862333"> <clone_1lib="NCI_CGAP_D10"> <tissue_type="Lung_Focal_Fibrosis"> <dev_stage="Adult"> <lab_host="DH10B (Life Technologies)"> <note="organ: Lung; vector: pT73-Pac (Pharmacia) with a modified polylinker; Site_1: Eco R I; Site_2: Not I;"> NCI_CGAP_D10 is a cDNA library containing the following tissue(s): A pool of lung Focal Fibrosis. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an Eco R I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first strand cDNA contains a library tag sequence that is located between the Not I site and the ATAGGGTC. (dP18 tail). The sequence tag for this library is TAG LIB-D10-H-D10 TAG TISSUE=lung with fibrosis TAG SEQ=ATACGGGC"	
BASE COUNT	130 a 182 c 242 g 137 t 1 others	
ORIGIN	Query Match 45.3%; Score 349.6; DB 14; Length 692; Site_1: Small; Site_2: SmtI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."	
	395 GGGTCACTGGCTGGGGGAGTGGACATGTAGAGCCTCCACCCCG 454 Db 692 GGGTCACTGGCTGGGGGAGTGGACATGTAGAGCCTCCACCCCG 633	
ORIGIN	Query Match 45.3%; Score 349.6; DB 14; Length 692; Site_1: Small; Site_2: SmtI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."	
	455 AGCAAGTGAAGGCTCCATAATGGAAACACATTGTGACCCAAATA 514 Db 632 AGCAAGTGA-GTCCCATAATGGAAACACATTGTGACCCAAATA 574	
ORIGIN	Query Match 45.3%; Score 349.6; DB 14; Length 692; Site_1: Small; Site_2: SmtI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."	
	515 OCTACAGGGAGACGAGCTGGCATGTGCGTGGAGCACATGTGCTGTGCGGGAGACACCC 574 Db 573 CNTACAGGGAGACGAGCTGGCATGTGCGTGGAGCACATGTGCGGGAGACACCC 514	
ORIGIN	Query Match 45.3%; Score 349.6; DB 14; Length 692; Site_1: Small; Site_2: SmtI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."	
	575 GGAGGGACTCATGGCAGGGCGACGCCGGGCGCTSTGTGAAAGGAGATGGCACCT 634 Db 513 GGAGGGACTCATGGCAGGGCGACGCCGGGCGCTGGTGTGCAAGGGATGGCACCT 454	
ORIGIN	Query Match 45.3%; Score 349.6; DB 14; Length 692; Site_1: Small; Site_2: SmtI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."	

Db 453 GGCTACAGGGGGCTGGTAGCTGGGAGGGCTGCCAACGGCTGGCA 394  
 Qy 695 TCTACACCCTGTACTACTACTTGACTGGATCCACCATGTCCTCAAAGCCGT 754  
 Db 393 TCTACACCCTGTACTACTACTTGACTGGATCCACCATGTCCTCAAAGCCGT 334  
 Qy 755 GAAGGGCC 763  
 Db 333 GAGTCAGGC 325

RESULT 14  
 LOCUS BQ003361/C  
 DEFINITION UT-H-ELL-azd-d-03-0-UT-s1 NCI\_CGAP\_ELL Homo sapiens linear EST 26-MAR-2002  
 ACCESSION BQ003361  
 VERSION BQ003361.1 GI:19728261  
 SOURCE human.  
 ORGANISM Homo sapiens  
 REFERENCE Mamalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 TITLE Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Straubberg, Ph.D.  
 Email: cgaps-r@mail.nih.gov  
 PUBLISHER Issue Procurement: Dr. Jose Mercuende  
 POLYA=Yes.

FEATURES source  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:584122"  
 /clone\_lib="NCI\_CGAP\_ELL"  
 /tissue\_type="Chondrosarcoma"  
 /dev\_stage="adult"  
 /lab\_host="DHLB (Life Technologies)"  
 /note="Organ: Left pelvis; Vector: pPTM3-Pac (Pharmacia)  
 NCI\_CGAP\_ELL is a normalized cDNA library containing the  
 following tissue(s): Chondrosarcoma. The library was  
 constructed according to Bonaldo, Lennon and Soares,  
 Genome Research, 6:791-806, 1996. First strand cDNA  
 synthesis was primed with an oligo-dT primer containing a  
 Not I site. Double stranded cDNA was ligated to an EcoRI  
 adaptor, digested with Not I, and cloned directionally  
 into pPTM3-Pac vector. The oligonucleotide used to prime  
 the synthesis of first-strand cDNA contains a library tag  
 sequence that is located between the Not I site and the  
 (GTT)18 tail. The sequence tag for this library is  
 ACATTGTCAC.  
 TAG\_LIBUT\_H\_ELL  
 TAG\_TISSUE="chondrosarcoma"  
 TAG\_SEQ="ACATTGTCAC"  
 BASE COUNT 130 a 176 c 242 g 140 t  
 ORIGIN

Db 688 ACTGGCTGGGGGTGGACATGATGAGCCCTCCACGCCATTCCCTGAAGCAG 629  
 Qy 460 GTGAGGGCCCATAAATGAAACACATTTGTCAGGAAATACCAACTT3CGCCCTAC 519  
 Db 628 GTGAGGTCCCATATGGAAACACATTTGTCAGCAGACATGCTGTGTCGGGAGG 579  
 Qy 520 ACGGAGACGAGCTGGCATCGTCGTCAGCAGTCGTCGTCGGGAAACACCGGAGG 579  
 Db 568 ACGGGAGACGAGCTGGCATCGTCGTCAGCAGACATGCTGTGTCGGGAGG 509

RESULT 15  
 LOCUS BF359130/C  
 DEFINITION BQ3-ET0063-090800-291-f02 ET0063 Homo sapiens mRNA linear EST 22-Nov-2000  
 ACCESSION BQ359130  
 VERSION 1  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 REFERENCE  
 AUTHORS Dias Neto, E., Garcia Correa, R., Veriovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W.J.R., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.  
 TITLE JOURNAL MEDLINE  
 COMMENT shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 20020663  
 Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil.  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: astimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
 (<http://www.ludwig.org.br/scripts/gethtml2.pi?pi1=QV3&t2=QV3-ET0063-00800-291-f02&t=2008-09-04-11>)  
 Seq primer: puc18 forward  
 High quality sequence start: 12  
 High quality sequence stop: 369.  
 FEATURES source  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="ET0063"  
 /dev\_stage="adult"  
 /note="Organ: lung-tumor; Vector: puc18; site\_1: small;  
 site\_2: small; A mini-library was made by cloning products

Qy 400 ACTGGCTGGGGATGGACACATGATGAGGCCATTCTCTGAAGCAG 459

derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pRC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT  
ORIGIN  
63 a 116 c 119 g 73 t

Query Match	Best Local Similarity	Score	DB	Length
348 CACCCCTGCCCTGCGCTCAGAGACCTTCCCCGGATGCCCTGCTGGTCACTGGTG	96.3%	334.2;	12;	371;
Db 371 CATGGTSCCCCCCTGCCTGGACCTTCCCCGGATGCCCTGCTGGTCACTGGTG	96.3%	Pred. No.	9.8e-61;	Mismatches 13;
Qy 408 GGCGGATGTGGACAAATGATGAGGGCCTCCACCGCACATTCCCTGAAGCAGGTGAAGGT	96.3%	No.	0;	Indels 0;
Db 311 GGGCATGTGGACAAATGATGAGGGCCTCCACCGCACATTCCCTGAAGCAGGTGAAGGT	96.3%			Gaps 0;
Qy 468 CCCCATTAATGGAAACCAATTTGAGCAAATACCCATTGGCGCCTAACGGAGA	96.3%			
Db 251 CCCCATTAATGGAAACCAATTTGAGCAAATACCCATTGGCGCCTAACGGAGA	96.3%			
Qy 528 CGAGCTCGCATGGCTGACGACATGCTGTCGGGACACCCGGAGGGACCTATG	96.3%			
Db 191 CGACGTCCGCATCATCCGTGACGACATGCTGTCGGGACAGCCAGAGGGACTCTG	96.3%			
Qy 588 CCAGGGCAGCTGGAGGGCCCTGCTGTCAGGTGATGGACCTGGCTGGAGGGGG	96.3%			
Db 131 CAAGGGGACTCTGGAGGGCCCTGCTGTCAGGTGATGGACCTGGCTGGAGGGGG	96.3%			
Qy 648 CGTGTCAAGCTGGGGCGAGGCCAACGGGCTGGCATCTACACC	96.3%			
Db 71 CGTGTCACTGGGACGGGGCTGTGGCCACCGGCCATGGCATCTACACC	96.3%			

Search completed: January 31, 2003, 06:50:06  
Job time : 2242 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

Om nucleic - nucleic search, using sw model  
Run on: January 31, 2003, 03:49:17 ; Search time 264 Seconds  
(without alignments)  
6576.859 Million cell updates/sec

Title: US-09-598-982-20  
Perfect score: 771

Sequence: 1 999ccccctcgagaaaaaat.....cgtgaaggccgcgcgtcgt 771

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Total number of hits satisfying chosen parameters:  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_101002;\*

1: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1980.DAT;\*  
2: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1981.DAT;\*  
3: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1982.DAT;\*  
4: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1983.DAT;\*  
5: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1984.DAT;\*  
6: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1985.DAT;\*  
7: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1986.DAT;\*  
8: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1988.DAT;\*  
9: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1989.DAT;\*  
10: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1990.DAT;\*  
11: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1991.DAT;\*  
12: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1992.DAT;\*  
13: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1993.DAT;\*  
14: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1994.DAT;\*  
15: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1995.DAT;\*  
16: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1996.DAT;\*  
17: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1997.DAT;\*  
18: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1998.DAT;\*  
19: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1999.DAT;\*  
20: /SIDS2/gcadata/geneseq/geneseq-emb1/NA2000.DAT;\*  
21: /SIDS2/gcadata/geneseq/geneseq-emb1/NA2001A.DAT;\*  
22: /SIDS2/gcadata/geneseq/geneseq-emb1/NA2001B.DAT;\*  
23: /SIDS2/gcadata/geneseq/geneseq-emb1/NA2002.DAT;\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

\* SUMMARIES

Result No.	Score	Query	Length	DB ID	Description
1	771	100.0	771	24	AS20775 DNA encoding human
2	769.4	99.8	771	24	AS20783 DNA encoding human
3	764.6	99.2	771	21	AS20175 Human beta-tryptase
4	764.6	99.2	771	24	AS20765 DNA encoding recom
5	763	99.0	771	24	AS20763 DNA encoding human
6	761.4	98.8	771	24	AS20776 DNA encoding human
7	759.8	98.5	771	24	AS20784 DNA encoding human
8	755	97.9	771	24	AS20777 DNA encoding human
9	755	97.9	771	24	AS20778 DNA encoding human

RESULT 1  
ID AAS20775 standard; DNA; 771 BP.  
XX AAS20775;  
XX DT 09-APR-2002 (first entry)  
XX DNA encoding human beta-II tryptase active site mutant H44A #1.  
XX Human; proteolytic tryptase; protease; recombinant beta-II tryptase;  
XX enzyme; mutant; ds.  
XX OS Homo sapiens.  
OS Synthetic.  
XX WO200198470-A2.  
XX PD 27-BBC-2001.  
XX 20-JUN-2001; 2001WO-US19681.  
XX PR 21-JUN-2000; 2000US-059982.  
XX PA (PROM-) PROMEGA CORP.  
XX Maffit M, Niles AL, Haak-Frendscho M;  
XX DR WPI: 2002-114578/15.  
XX P-PSDB; AU12009.  
XX PT DNA construct for producing enzymatically-inactive proteolytic



QY	61	CAGGTTGAGCTTGTGTCACCTTCGGGGCTCCATC	ACTTGATGCACTTCGGGGCTCCATC	120	XX	(PROM') PROMEGA CORP.
Db	61	CAGGTGAGGCTGTGAGATGCCAAGGCCATACTGGATGCACTTCGGGGCTCCATC	120	PA		
PI				XX		
QY	121	CACCCCGAGTGGGCAACTGGGGGACGACCTCTACTACCCGCCAGATGGGACGATCTGGTC	180	DR	WPI; 2000-053300/04.	
Db	121	CACCCCGAGTGGGCAACTGGGGGACGACCTCTACTACCCGCCAGATGGGACGATCTGGTC	180	DR	P-PSB; AAY55011.	
QY	241	ACCAGATCTTGCACCCACAGTCTAACCGCCAGATGGGACGGGCTCAGGATCTGGC	300	XX		
Db	241	ACCAGATCTTGCACCCACAGTCTAACCGCCAGATGGGACGGGCTCAGGATCTGGC	300	PT	New DNA expression construct for production of enzymatically active recombinant human beta-trypsinase.	
QY	181	GCCCCTAGGGTCACTGGGGGACGACCTCTACTACCCGCCAGATGGGACGATCTGGTC	240	XX		
Db	181	GCCCCTAGGGTCACTGGGGGACGACCTCTACTACCCGCCAGATGGGACGATCTGGTC	240	PS	Disclosure; Page 43-44; 50pp; English.	
QY	301	CTGGACTGGAGGCCGGTGAACCTGCGGACGGCTCAGGATCTGGC	360	XX		
Db	301	CTGGACTGGAGGCCGGTGAACCTGCGGACGGCTCAGGATCTGGC	360	CC	This sequence encodes the human beta-trypsinase. The invention relates to a DNA expression construct comprising (5' to 3') a promoter linked to a signal sequence which is linked to a sequence encoding human beta-trypsinase. The DNA construct is useful for transforming host cells express, post translationaly process and secrete enzymatically active human trypsinase. The method is useful for the production of large amount of trypsinase with defined specifications. The transformant is useful for pharmaceutical studies, combinatorial library screens and X-ray crystallographic studies. The trypsinase produced allows for the development of trypsinase agonists and/or antagonists, is useful as an antigen to generate human trypsinase antibodies in various animals, can be used in screening for compounds which act as trypsinase inhibitors, antagonists, agonists etc. and to assay for the presence of trypsinase in biological or other solutions. Trypsinase inhibitors, antagonists, agonists may be useful as therapeutics. The trypsinase does not require any post-expression or post-purification modifications or manipulations to initiate trypsinase activity and it has enzymatic activity which compares favourably with cadheric trypsinase. The availability of enzymatically active trypsinase facilitates the large scale screening of combinatorial libraries for specific trypsinase inhibitors as potential therapeutics advances the understanding of the biological significance of trypsinase in mast cell mediated diseases. The trypsinase can be used to detect low levels of trypsinase.	
QY	421	ATATGAGGGCTCCACGGCATTCCTGAGCAGGTGAAGGTCCATATGGAA	480	CC		
Db	421	ATATGAGGGCTCCACGGCATTCCTGAGCAGGTGAAGGTCCATATGGAA	480	CC		
QY	361	GCCTCTAGAGACCTTCCCCTGGGATGCGGTGCGGGTACTGCTGGGATGCGGCT	300	CC		
Db	361	GCCTCTAGAGACCTTCCCCTGGGATGCGGTGCGGGTACTGCTGGGATGCGGCT	300	CC		
QY	301	CTGGACTGGAGGCCGGTGAACCTGCGGACGGCTCAGGATCTGGC	360	CC		
Db	301	CTGGACTGGAGGCCGGTGAACCTGCGGACGGCTCAGGATCTGGC	360	CC		
QY	421	ATATGAGGGCTCCACGGCATTCCTGAGCAGGTGAAGGTCCATATGGAA	480	CC		
Db	421	ATATGAGGGCTCCACGGCATTCCTGAGCAGGTGAAGGTCCATATGGAA	480	CC		
QY	481	AACCATATGGTGGACCAAATAACACCTTGGGCTTACAGGGAGAACGAGTCGCACT	540	CC		
Db	481	AACCATATGGTGGACCAAATAACACCTTGGGCTTACAGGGAGAACGAGTCGCACT	540	CC		
QY	541	GTCCGTGACCAACGACGTCGTTGGGGAAACCCGGGACTCATGGCAGGGACTCC	600	CC		
Db	541	GTCCGTGACCAACGACGTCGTTGGGGAAACCCGGGACTCATGGCAGGGACTCC	600	CC		
QY	601	GGAGGGCCCTGGTGTGCAAGGTGATGGGACCTGGAGGGGGGGGGTCACTGG	660	CC		
Db	601	GGAGGGCCCTGGTGTGCAAGGTGATGGGACCTGGAGGGGGGGGGTCACTGG	660	CC		
QY	661	GGCGAGGGCTGTGCCAGCCAAACGGGCTGGCARCTACCCCGGTACTACTTG	720	CC		
Db	661	GGCGAGGGCTGTGCCAGCCAAACGGGCTGGCARCTACACCCGGTACTACTTG	720	CC		
QY	721	GACTGATCCACACTATGTCCCCAAAAGCGTGAAGCGGCCGCGTGTCT	771	CC		
Db	721	GACTGATCCACACTATGTCCCCAAAAGCGTGAAGCGGCCGCGTGTCT	771	CC		
RESULT 3						
RAZ40175	ID	AA240175 standard; DNA; 771 BP.	SO	Sequence 771 BP: 149 A; 256 C; 242 G; 124 T; 0 other;		
RAZ40175;	DT	18-FEB-2000 ( first entry)	Query Match	99.2%; Score 764.6; DB 21; Length 771;		
RAZ40175;	XX	Human beta-trypsinase coding sequence.	Best Local Similarity	99.5%; Pred. No. 5e-147;		
RAZ40175;	XX	Beta-trypsinase; human; DNA expression construct; protein production; antibody generation; ss. Homo sapiens. WO960139-A1.	Matches	767; Conservative 0; Mismatches 4; Indels 0; Gaps		
RAZ40175;	XX		QY	1 GGGCCCTCCGAGAAAGAATCTGGGGGTCAGGAGGCCCGAGGCAAGTGGCCCTGG 60		
RAZ40175;	XX		Db	1 GGGCCCTCCGAGAAAGAATCTGGGGGTCAGGAGGCCCGAGGCAAGTGGCCCTGG 60		
RAZ40175;	XX		QY	61 CAGGTCAGGCTGAGACTCCAGGCCCATACTGGATGCACTCTGGGGCTCCCTCATC 120		
RAZ40175;	XX		Db	61 CAGGTCAGGCTGAGACTCCAGGCCCATACTGGATGCACTCTGGGGCTCCCTCATC 120		
RAZ40175;	XX		QY	121 CACCCCGAGTGGTGTGACCGCCAGGCCACTGGGGACGGCTCAGGATCTGGC 180		
RAZ40175;	XX		Db	121 CACCCCGAGTGGTGTGACCGCCAGGCCACTGGGGACGGCTCAGGATCTGGC 180		
RAZ40175;	XX		QY	181 GCCCTCAGGGTCAACTGGGGACGACCTCTACTACCCGCCAGATGGGACGATCTGGTC 240		
RAZ40175;	XX		Db	181 GCCCTCAGGGTCAACTGGGGACGACCTCTACTACCCGCCAGATGGGACGATCTGGTC 240		
RAZ40175;	XX		QY	241 AGCAGATCATCTGGGCCACAGTCTAGACCGCCAGATCGGAGGGACATGCCCTG 300		
RAZ40175;	XX		Db	241 AGCAGATCATCTGGGCCACAGTCTAGACCGCCAGATCGGAGGGACATGCCCTG 300		
RAZ40175;	XX		QY	301 CTGGACTGGAGGCCGGTGAACCTGCGGACGGCTCAGGATCTGGGACGTCACCTGGCCCT 360		
RAZ40175;	XX		Db	301 CTGGACTGGAGGCCGGTGAACCTGCGGACGGCTCAGGATCTGGGACGTCACCTGGCCCT 360		
RAZ40175;	XX		QY	361 GCCTAGAGACCTTCCCCTGGGATGCGGTACTGGGGCGAGATGGAC 420		
RAZ40175;	XX		Db	361 GCCTAGAGACCTTCCCCTGGGATGCGGTACTGGGGCGAGATGGAC 420		
RAZ40175;	XX		QY	421 ATATGAGGGCTCCACGGCATTCCTGAGCAGGTGAAGGTCCATATGGAA 480		

Db 421 AATGATGAGGCCCTCCACCGCCATTCTCTGAAAGCAGGTGAGGTCCTAATGCAA 480  
 QY 481 AACCACATTTGAGGCCAAATACCGACTTGGCCCATCACGGAGAGACGGTCCGATC 540  
 Db 481 AACCACATTTGAGGCCAAATACCGACTTGGCCCATCACGGAGAGACGGTCCGATC 540  
 QY 541 GTCGGTGGACATCGTGTGGGGACACCCGGGGACTCATGCCAGGGACTCC 600  
 Db 541 GTCGGTGGACATCGTGTGGGGACACCCGGGGACTCATGCCAGGGACTCC 600  
 QY 601 GGAGGGCCCTGGTGGAAGGTAATGCCACCTGCGAGGGGGCTGGTCACTGG 660  
 Db 601 GGAGGGCCCTGGTGGAAGGTAATGCCACCTGCGAGGGGGCTGGTCACTGG 660  
 QY 661 GGGGAGGGCTGTGTCGCCAACCGGCCATCACCCGGTGTACCTACTACTTG 720  
 Db 661 GGGGAGGGCTGTGTCGCCAACCGGCCATCACCCGGTGTACCTACTACTTG 720  
 QY 721 GACTGGATCCACCACTATGTCCTCAAAGGCCGAGCGGGCGCGTGT 771  
 Db 721 GACTGGATCCACCACTATGTCCTCAAAGGCCGAGCGGGCGCGTGT 771

RESULT 4  
 AA\$20765  
 ID AA\$20765 standard; DNA: 771 BP.  
 AC XX  
 XX XX  
 DT 09-APR-2002 (first entry)  
 XX DE DNA encoding recombinant human beta-II tryptase.  
 KW Human; proteolytic tryptase; protease; recombinant beta-II tryptase;  
 KW enzyme; gene; mutant; ds.  
 XX OS Homo sapiens.  
 OS Synthetic.  
 XX FH Key  
 FT 7..756 Location/Qualifiers  
 FT /\*tag= a /partial /product= "Recombinant beta-II tryptase"  
 FT /note= "This sequence lacks a start codon"  
 XX PN WO200198470-A2.  
 XX PD 27-DEC-2001.  
 XX P F 20-JUN-2001; 2001WO-US19681.  
 XX PR 21-JUN-2000; 2000US-0598982.  
 XX PA (PROM-) PROMEGA CORP.  
 XX Maffit M, Niles AL, Haak-Prendscho M;  
 DR WPI: 2002-114578/15.  
 DR P-PSDB; AU12007.  
 XX PT DNA construct for producing enzymatically-inactive proteolytic  
 PT triptase, comprises DNA sequence encoding proteolytic triptase having  
 an active site mutation -  
 XX PS Claim 43; Page 75-77; 126pp; English.  
 XX  
 CC The present invention relates to recombinant human proteolytic  
 CC triptases, active site mutants of these triptases and the methods for  
 CC producing these. The method involves the production of a DNA expression  
 CC construct comprising a promoter operably linked to a secretion signal  
 CC sequence which is operably linked to a DNA sequence encoding a  
 CC proteolytic triptase with an active site mutation (the construct drives

CC expression of a mature proteolytic triptase that lacks enzymatic activity  
 CC due to the active site mutation, in hosts transformed to contain the  
 CC construct). The method is useful for producing enzymatically-active  
 CC beta-II triptase. The active site mutants of proteolytic triptase provide  
 CC a tool to investigate the structural and functional properties of the  
 CC protease and its enzymatic activity, and for modelling studies. The  
 CC enzymatically-active, recombinant proteolytic triptase produced are  
 CC useful as an antigen to generate anti-human triptase antibodies  
 CC and in drug screening for compounds which act as triptase inhibitors,  
 CC antagonists, agonists, etc. The present sequence encodes for recombinant  
 CC human beta-II triptase.

XX SQ Sequence 771 BP; 149 A; 256 C; 242 G; 124 T; 0 other;

Query Match 99.2%; Score 744.6; DB 24; Length 771;  
 Best Local Similarity 99.5%; Pred. No. 5e-147; Matches 0; Mismatches 4; Indels 0; Gaps 0;

Matches 767; Conservative 0; Gaps 0;

Db 1 GGGCCCTCTGAGAAAGATCCGGGGGTAGAGGCCCAAGGGCAAGGCCCTGG 60  
 Db 1 GGGCCCTCTGAGAAAGATCCGGGGGTAGAGGCCCAAGGGCAAGGCCCTGG 60

QY 61 CAGGTGAGGCTGAGAGTCACCCCATACTGGTACTGACTTGCAGGGCTTCATC 120  
 Db 61 CAGGTGAGGCTGAGAGTCACCCCATACTGGTACTGACTTGCAGGGCTTCATC 120  
 QY 121 CACCCCCAGTGGGTGTCACCCCGCCGGGGACTGGTGGGACCGTCAGGATCTGGC 180  
 Db 121 CACCCCCAGTGGGTGTCACCCCGCCGGGGACTGGTGGGACCGTCAGGATCTGGC 180

QY 181 GCCTCTCAGGGTCACTGGGGAGCACCTACTACCAAGGACAGCTGCTGGC 180  
 Db 181 GCCTCTCAGGGTCACTGGGGAGCACCTACTACCAAGGACAGCTGCTGGC 180

QY 241 AGCAGGATCATCGTGCACCCAGGTCTACACGCCAGATGGAGGGATTCVCGCCTG 300  
 Db 241 AGCAGGATCATCGTGCACCCAGGTCTACACGCCAGATGGAGGGATTCVCGCCTG 300  
 QY 301 CTGGAGCTGGAGGCCGGTGAAGGTCTCCACCGCAGATGGAGGGATTCVCGCCTG 360  
 Db 361 GCCTCTCAGGGTCACTGGGGAGCACCTACTACCAAGGACAGCTGCTGGC 360

QY 361 GCCTCTCAGGGTCACTGGGGAGCACCTACTACCAAGGACAGCTGCTGGC 360  
 Db 361 GCCTCTCAGGGTCACTGGGGAGCACCTACTACCAAGGACAGCTGCTGGC 360

QY 421 AATGATGAGGCCCTCCACCGCCATTCTCTGAGAGGTCAAGGTCCTAATGAA 480  
 Db 421 AATGATGAGGCCCTCCACCGCCATTCTCTGAGAGGTCAAGGTCCTAATGAA 480  
 QY 481 ACCACATTTGAGGCCAAATACCGACTTGGCCATCACGGAGAGACGGTCCGATC 540  
 Db 481 ACCACATTTGAGGCCAAATACCGACTTGGCCATCACGGAGAGACGGTCCGATC 540  
 QY 541 GTCGGTGGACATCGTGTGGGGACACCCGGGGACTCATGCCAGGGACTCC 600  
 Db 541 GTCGGTGGACATCGTGTGGGGACACCCGGGGACTCATGCCAGGGACTCC 600  
 QY 601 GAGGGCCCTGGTGGAAGGTAATGCCACCTGGCCATCACGGAGAGACGGTCCGATC 660  
 Db 601 GAGGGCCCTGGTGGAAGGTAATGCCACCTGGCCATCACGGAGAGACGGTCCGATC 660  
 QY 661 GGGGAGGGCTGTGTCGCCAACCGGCCATCACCCGGTGTACCTACTACTTG 720  
 Db 661 GGGGAGGGCTGTGTCGCCAACCGGCCATCACCCGGTGTACCTACTACTTG 720

RESULT 5  
 AA\$20763

ID AAS20763 Standard; DNA; 771 BP.  
 XX  
 AC AAS-20763;  
 XX  
 DT 09-APR-2002 (first entry)  
 XX  
 DE DNA encoding human beta-I trypsinase.  
 XX  
 KW Human; proteolytic trypsinase; protease; beta-I trypsinase;  
 KW enzyme; gene; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 7.. 756  
 FT /\*tag= a  
 FT Partial /product= "Beta-I trypsinase"  
 FT /note= "This sequence lacks a start codon"  
 XX  
 PF 20-JUN-2001; 2001WO-US19681.  
 PN WO200198470-A2.  
 PR 21-JUN-2000; 2000US-0598982.  
 PD 27-DEC-2001.  
 XX  
 PA (PROM-) PROMEGA CORP.  
 XX  
 PT Maffit M, Niles AL, Haak-Frendscho M;  
 XX  
 DR WPI; 2002-114578/15.  
 XX  
 PS Disclosure; Page 72-74; 126pp; English.  
 XX  
 CC The present invention relates to recombinant human proteolytic trypsinases, active site mutants of these trypsinases and the methods for producing these. The method involves the production of a DNA expression construct comprising a promoter operably linked to a secretion signal sequence which is operably linked to a DNA sequence encoding a proteolytic trypsinase with an active site mutation (the construct drives expression of a mature proteolytic trypsinase that lacks enzymatic activity due to the active site mutation in hosts transformed to contain the construct). The method is useful for producing enzymatically-active beta-II trypsin. The active site mutants of proteolytic trypsinase provide a tool to investigate the structural and functional properties of the protease and its enzymatic activity, and for modeling studies. The enzymatically-active, recombinant proteolytic trypsinase produced are useful as an antigen to generate anti-human trypsinase antibodies and in drug screening for compounds which act as trypsinase inhibitors, antagonists, agonists, etc. The present sequence encodes for human beta-I trypsinase.  
 XX  
 SQ Sequence 771 BP; 149 A; 257 C; 241 G; 124 T; 0 other;  
 Query Match 99.0%; Score 763; DB 24; Length 771;  
 Best Local Similarity 99.4%; Pred. No. 1.1e-146;  
 Matches 766; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 GGCCCCCTGAGAAAAGAATCGTGGGGCTCAGAGGCCCAAGGAGCAAGTGCCCG 60  
 DB 1 GGGCCCTGAGAAAAGAATCGTGGGGCTCAGAGGCCCAAGGAGCAAGTGCCCG 60  
 QY 61 CAGGTGAGCTGAGTCACGGCCCATACTGATGACTTCGCGGGGCTCCCTAC 120  
 DB 61 CAGGTGAGCTGAGTCACGGCCCATACTGAGTCACGGCC 120  
 OQ 121 CACCCCCAGTGGGCTCTGACCCGGCCGGGTGCTGGGACGGGACTGCAAGATCTGCC 180  
 DB 121 CACCCCCAGTGGGCTGACCCGGCAAGCAGCTGTTGGACCGGAAGTCAGATCTGCC 180  
 OQ 181 GCCCTCAGGGTCACTGGGAGCACCTACTACAGGACCACTGCGGCCGTC 240  
 DB 181 GCCCTCAGGGTCACTGGGAGCACCTACTACAGGACCACTGCGGCCGTC 240  
 OQ 241 AGCAGGATCATCTGCACCCACAGTCTACACGCCAGATCGAGCGACATCGCCCTG 300  
 DB 241 AGCAGGATCATCTGCACCCACAGTCTACACGCCAGATCGAGCGACATCGCCCTG 300  
 OQ 301 CTGGAGCTGGAGGCCGGTCAAAGGTCTCCAGGTCACGGTACCCGGCCCT 360  
 DB 301 CTGGAGCTGGAGGCCGGTCAAAGGTCTCCAGGTCACGGTACCCGGCCCT 360  
 OQ 361 GCCCTCAGAGACACTGCCCGGGATGCGGTGCTGGGACTCTGGGGATGTTGG 420  
 DB 361 GCCCTCAGAGACCTGCCCGGGATGCGGTGCTGGGACTCTGGGGATGTTGG 420  
 OQ 421 ATATGATGAGGCCCTCCACGCCATTCTCTGAGCAGGTGAGGTCCTCATATGGAA 480  
 DB 421 ATATGATGAGGCCCTCCACGCCATTCTCTGAGCAGGTGAGGTCCTCATATGGAA 480  
 OQ 481 ACCACATTTGTGAGCATACTCTGCTGGGGATGGCTCCATGGGACATCC 540  
 DB 481 ACCACATTTGTGAGCATAATTACCCATGGGGCTACACGGAGACGTCGCACTC 540  
 OQ 541 GTCCGTGAGCAGATGCTGTCGCGGGAAACCCGGGAGGACTATGCCAGGSGGACTCC 600  
 DB 541 GTCCGTGAGCAGATGCTGTCGCGGGAAACCCGGGAGGACTATGCCAGGSGGACTCC 600  
 OQ 601 GGAGGGCCCTGGTGTGCAAGGTGAATSGCACCTGGCTGCAAGGGGGGTGTCAGCTG 660  
 DB 601 GGAGGGCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAAGGGGGGTGTCAGCTG 660  
 OQ 721 GACTGGATCACCACATGCCCCAAAAGCCGTGAAGGGGGGGGTGT 771  
 DB 721 GACTGGATCACCACATGCCCCAAAAGCCGTGAAGGGGGGGGTGT 771  
 RESULT 6  
 AAS2076 ID AAS2076 standard; DNA; 771 BP.  
 XX  
 AC AAS2076;  
 XX  
 DT 09-APR-2002 (first entry)  
 XX  
 DE DNA encoding human beta-II trypsinase active site mutant D91A #1.  
 XX  
 KW Human; proteolytic trypsinase; protease; recombinant beta-II trypsinase;  
 KW enzyme; mutant; ds.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN WO200198470-A2.  
 XX  
 PD 27-DEC-2001.  
 XX  
 PF 20-JUN-2001; 2001WO-US19681.  
 XX  
 PR 21-JUN-2000; 2000US-0598982.  
 XX  
 PA (PROM-) PROMEGA CORP.  
 XX  
 PI Maffit M, Niles AL, Haak-Frendscho M;

DR WPI; 2002-114578/15.  
 DR P-PSDB; AAU12010.  
 XX  
 PT DNA construct for producing enzymatically-inactive proteolytic trypsinase, comprises DNA sequence encoding proteolytic trypsinase having an active site mutation -  
 PT XX  
 PS Claim 7; Page 86-87; 126pp; English.

The present invention relates to recombinant human proteolytic trypsinases, active site mutants of these trypsinases and the methods for producing these. The method involves the production of a DNA expression construct comprising a promoter operably linked to a secretion signal sequence which is operably linked to a DNA sequence encoding a proteolytic trypsinase with an active site mutation (the construct drives expression of a mature proteolytic trypsinase that lacks enzymatic activity due to the active site mutation, in hosts transformed to contain the construct). The method is useful for producing enzymatically-active beta-II trypsinase. The active site mutants of proteolytic trypsinase provide a tool to investigate the structural and functional properties of the protease and its enzymatic activity, and for modelling studies. The enzymatically-active, recombinant proteolytic trypsinase produced are useful as an antigen to generate anti-human trypsinase antibodies and in drug screening for compounds which act as trypsinase inhibitors, antagonists, agonists, etc. AAS2075-AAS20790 encode for recombinant human beta-II trypsinase active site mutants.

Sequence 771 BP; 149 A; 256 C; 242 G; 124 T; 0 other;

Query Match 98.8%; Score 761.4; DB 24; Length 771;  
 Best Local Similarity 99.2%; Pred. No. 2.2e-146;  
 Matches 765; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 GGGCCCTGAGAAAGAAGTGTGGGGTCAGAGGGCCCCAGAGCAGTGCCTCG 60  
 Db 1 GGGCCCTGAGAAAGAAGTGTGGGGTCAGAGGGCCCCAGAGCAGTGCCTCG 60

Qy 61 CAGGGAGCTGAGATCTACGGCCCATACTGGATCACTCTGGGGCTCCATC 120  
 Db 61 CAGGGAGCTGAGATCTACGGCCCATACTGGATCACTCTGGGGCTCCATC 120

Qy 121 CACCCCACTGGGTCTGACGCCCGGGCTGGTGGGACGGAGCTCAGGATCTGGC 180  
 Db 121 CACCCCACTGGGTCTGACGCCCGGGCTGGTGGGACGGAGCTCAGGATCTGGC 180

Qy 181 GGCCTCAGGGTGCACTGGGGGAGAGCACCTCTACAGGACGAGTCAGGATCTGGC 240  
 Db 181 GGCCTCAGGGTGCACTGGGGGAGAGCACCTCTACAGGACGAGTCAGGATCTGGC 240

Qy 241 AGCGAGTATCGTCACCCACAGTCTACACCCAGATCGGAAGCCTAACGCCCTG 300  
 Db 241 AGCAGAGTATCGTCACCCACAGTCTACACCCAGATCGGAAGCCTAACGCCCTG 300

Qy 301 CTGGAGCTGGAGACCCGGTAAGCTCTACGCCAGTCACAGGTACCCCTGCCCT 360  
 Db 301 CTGGAGCTGGAGACCCGGTAAGCTCTACGCCAGTCACAGGTACCCCTGCCCT 360

Qy 361 GCCTCTAGAGACCTTCCCGGGGGCGCTGCTGGTCACTGGTGGGAGCTGGG 420  
 Db 361 GCCTCTAGAGACCTTCCCGGGGGCGCTGCTGGTCACTGGTGGGAGCTGGG 420

Qy 421 AATGATGAGACCTTCCCGGGGGCGCTGCTGGTCACTGGTGGGAGCTGGG 480  
 Db 421 AATGATGAGACCTTCCCGGGGGCGCTGCTGGTCACTGGTGGGAGCTGGG 480

Qy 481 AACACATTTGTGAGCAAATAACCACTTGGGGCTAACGGGGAGACGTGGCATC 540  
 Db 481 AACACATTTGTGAGCAAATAACCACTTGGGGCTAACGGGGAGACGTGGCATC 540

Qy 541 GTCCCGGACGACATCTGTGCGCCGAACACCCGGGACTCATGGCAGGGACTCC 600  
 Db 541 GTCCCGGACGACATCTGTGCGCCGAACACCCGGGACTCATGGCAGGGACTCC 600

Qy 601 GAGGGCCCTGGTGCAGTCATGGCACCTGGCTGCAGGGGGCTGGTCAGCTGG 660  
 Db 601 GGAGGCCCTGGTGCAGTCATGGCACCTGGCTGCAGGGGGCTGGTCAGCTGG 660  
 Qy 661 GGCAGGGCTGTGCCAGGCCAACGGCCTGCACTACACCGGTGACCTACTTG 720  
 Db 661 GCGAGGGCTGTGCCAGGCCAACGGCCTGCACTACACCGGTGACCTACTTG 720  
 Qy 721 GACTGGATCACCACATATGCCAAAAGCCGTAACSGGGGCCGRCGT 771  
 Db 721 GACTGGATCACCACATATGCCAAAAGCCGTAACSGGGGCCGRCGT 771

RESULT 7

AAS20784

AAS20784 standard; DNA; 771 BP.

XX

AC

AAS20784;

XX

DT

09-APR-2002 (first entry)

XX

DE

DNA encoding human beta-II trypsinase active site mutant D91A #3.

XX

KW

Human; proteolytic trypsinase; protease; recombinant beta-II trypsinase; enzyme; mutant; ds.

XX

OS

Synthetic.

XX

PN

W0200198470-A2.

XX

PD

27-DBC-2001.

XX

PP

20-JUN-2001; 2001WO-US19681.

XX

PR

21-JUN-2000; 2000US-0598982.

XX

PA

(PROM-) PROMEGA CORP.

XX

PI

Maffit M,

Niles AL,

Haak-Frendscho M;

XX

DR

WPI; 2002-114578/15.

XX

P-PSDB;

AAU12018.

XX

PB

2000198470-A2.

XX

PT

DNA construct for producing enzymatically-inactive proteolytic trypsinase having an active site mutation -

XX

PS

Claim 7; Page 107-109; 126pp; English.

The present invention relates to recombinant human proteolytic trypsinases, active site mutants of these trypsinases and the methods for producing these. The method involves the production of a DNA expression construct comprising a promoter operably linked to a DNA sequence encoding a proteolytic trypsinase with an active site mutation (the construct drives expression of a mature proteolytic trypsinase that lacks enzymatic activity due to the active site mutation, in hosts transformed to contain the construct). The method is useful for producing enzymatically-active beta-II trypsinase. The active site mutants of proteolytic trypsinase provide a tool to investigate the structural and functional properties of the protease and its enzymatic activity, and for modelling studies. The enzymatically-active, recombinant proteolytic trypsinase produced are useful as an antigen to generate anti-human trypsinase antibodies and in drug screening for compounds which act as trypsinase inhibitors, antagonists, agonists, etc. AAS2075-AAS20790 encode for recombinant human beta-II trypsinase active site mutants.

XX

Sequence 771 BP; 149 A; 257 C; 241 G; 124 T; 0 other;

SQ

Query Match

98.5%;

Score 759.8;

DB 24;

Length 771;

Best Local Similarity 99.1%;

Pred. No. 4.7e-146;

Matches 764;

Conservative 0;

Mismatches 7;

Indels 0;

Gaps 0;

QY	1	GGGCCCCCTGGGAAAGAATGTCGGGATCAGAGGCCAGGGAGTGGCCCTCG	PP	20-JUN-2001; 2001WO-US19681.
Db	1	GGGCCCCCTGGGAAAGAATGTCGGGATCAGAGGCCAGGGAGTGGCCCTCG	XX	
QY	61	CAGGTGAGCTGAGACTCCACGGCCATACTGGATGCACTCTGGGGGCTCCTCATC	PA	(PROM-) PROMEGA CORP.
Db	61	CAGGTGAGCTGAGACTCCACGGCCATACTGGATGCACTCTGGGGGCTCCTCATC	PT	Maffit M., Niles AL, Haak-Frednscho M;
QY	121	CACCCCACTGGGTTGCACTGGGGAGGAGCACTGGGGCTGGGGCTGGGGCTGG	XX	
Db	121	CACCCCACTGGGTTGCACTGGGGAGGAGCACTGGGGCTGGGGCTGGGGCTGG	DR	WPI; 2002-114578/15.
QY	181	GCCTCAGGCTGAGACTCCACGGCCATACTGGATGCACTCTGGGGGCTCCTCATC	XX	21-JUN-2000; 2000US-0598982.
Db	181	GCCTCAGGCTGAGACTCCACGGCCATACTGGATGCACTCTGGGGGCTCCTCATC	DR	P-PSDB; AAU12011.
QY	241	AGCAGATCATCGTGACCCACAGTCTACACCACGGCCAGTCGGGGCTGGGGCTGG	XX	DNA construct for producing enzymatically-inactive proteolytic trypsin
Db	241	AGCAGATCATCGTGACCCACAGTCTACACCACGGCCAGTCGGGGCTGGGGCTGG	PT	comprises DNA sequence encoding proteolytic trypsin having an active site mutation
QY	301	CTGGAGCTGGAGGCCGGGAAGGCTCCAGCCAGTCCACAGGTACCCCTGGCCCT	XX	Claim 7; Page 88-90; 126pp; English.
Db	301	CTGGAGCTGGAGGCCGGGAAGGCTCCAGCCAGTCCACAGGTACCCCTGGCCCT	PS	
QY	361	GCCTCAGGACTCCGG	XX	The present invention relates to recombinant human proteolytic trypsinases, active site mutants of these trypsinases and the methods for producing these. The method involves the production of a DNA expression construct comprising a promoter operably linked to a secretion signal sequence which is operably linked to a DNA sequence encoding a proteolytic trypsin with an active site mutation (the construct drives expression of a mature proteolytic trypsin that lacks enzymatic activity due to the active site mutation, in hosts transformed to contain the construct). The method is useful for producing enzymatically-active beta-II trypsin. The active site mutants of proteolytic trypsinase provide a tool to investigate the structural and functional properties of the protease and its enzymatic activity, and for modelling studies. The enzymatically-active recombinant proteolytic trypsin produced are useful as an antigen to generate anti-human trypsin antibodies, and in drug screening for compounds which act as trypsin inhibitors, antagonists, agonists, etc. AAS20775-AAS20790 encode for recombinant human beta-II trypsin active site mutants.
Db	361	GCCTCAGGACTCCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	CC	
QY	361	GCCTCAGGACTCCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	CC	
Db	361	GCCTCAGGACTCCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	CC	
QY	421	AATGATGAGGCGCTCCACGCCATTTCCTCTGAAGCAGGTGAATGTCCTGAA	XX	
Db	421	AATGATGAGGCGCTCCACGCCATTTCCTCTGAAGCAGGTGAATGTCCTGAA	CC	
QY	481	AACCACATTGTGAGCAAATAACCACTTGGCCTACACGGGAGACGCTCGCGCATC	XX	
Db	481	AACCACATTGTGAGCAAATAACCACTTGGCCTACACGGGAGACGCTCGCGCATC	CC	
QY	541	GTCCGTGACACATGTGTCGGGAAACCCGAGGACTCAGCCGGGCACTCC	XX	
Db	541	GTCCGTGACACATGTGTCGGGAAACCCGAGGACTCAGCCGGGCACTCC	CC	
QY	601	GGGAGGCCCTCTGGTGTGCAAGGTGATGGACCTGGGGGGGGGGGGGGGGGGGG	XX	
Db	601	GGGAGGCCCTCTGGTGTGCAAGGTGATGGACCTGGGGGGGGGGGGGGGGGG	CC	
QY	660	GG	XX	
Db	660	GG	CC	
QY	661	GG	XX	
Db	661	GG	CC	
QY	720	GG	XX	
Db	720	GG	CC	
QY	721	GACTGATCCACCATGTCGCCAAAACCGTGAAAGCCGGGAGGCTACTACTTG	XX	
Db	721	GACTGATCCACCATGTCGCCAAAACCGTGAAAGCCGGGAGGCTACTACTTG	CC	
QY	771	GACTGGATCCACCATGTCGCCAAAACCGTGAAAGCCGGGAGGCTACTACTTG	XX	
RESULT 8				
AAS2077				
ID				
AAS2077				
standard; DNA; 771 BP.				
XX				
AC				
AAS2077;				
09-APR-2002				
(first entry)				
DT				
XX				
DE				
DNA encoding human beta-II trypsin active site mutant S194A #1.				
XX				
KW				
Human; proteolytic tryptase; protease; recombinant beta-II trypsin;				
enzyme; mutant; ds.				
XX				
OS				
Homo sapiens.				
OS				
Synthetic.				
XX				
PN				
WO20198470-A2.				
PD				
27-DEC-2001.				

QY 481 AACCAACATTGTGAGCCAAATACCCACCTGGCCTACACGGAGACGTCGCATC 540  
 CC and in drug screening for compounds which act as trypsin inhibitors,  
 CC antagonists, agonists, etc. RAS2075\_RNS20790 encode for recombinant  
 CC human beta-II trypsin active site mutants.  
 XX Sequence 771 BP; 152 A; 255 C; 241 G; 123 T; 0 other;

QY 481 AACCAACATTGTGAGCCAAATACCCACCTGGCCTACACGGAGACGTCGCATC 540  
 CC and in drug screening for compounds which act as trypsin inhibitors,  
 CC antagonists, agonists, etc. RAS2075\_RNS20790 encode for recombinant  
 CC human beta-II trypsin active site mutants.  
 XX Sequence 771 BP; 152 A; 255 C; 241 G; 123 T; 0 other;

QY 541 GTCCGGTGAACGACANCGTGTGGGGAAACACCGGGAGGACTCATGGCGGCGACTC 600  
 DB 541 GTCCGGTGAACGACANCGTGTGGGGAAACACCGGGAGGACTCATGGCGGCGACTC 600  
 CC and in drug screening for compounds which act as trypsin inhibitors,  
 CC antagonists, agonists, etc. RAS2075\_RNS20790 encode for recombinant  
 CC human beta-II trypsin active site mutants.  
 XX Sequence 771 BP; 152 A; 255 C; 241 G; 123 T; 0 other;

QY 601 GGAGGCCCTGTTGGTGCAGGGATGACCGCTGGCTGAGGGGGTGCTAACGGCG 660  
 DB 601 GGAGGCCCTGTTGGTGCAGGGATGACCGCTGGCTGAGGGGGTGCTAACGGCG 660  
 CC and in drug screening for compounds which act as trypsin inhibitors,  
 CC antagonists, agonists, etc. RAS2075\_RNS20790 encode for recombinant  
 CC human beta-II trypsin active site mutants.  
 XX Sequence 771 BP; 152 A; 255 C; 241 G; 123 T; 0 other;

QY 661 GCGGAGGGCTGTGCCAGCCACGGGCTGGACATCACCGGGTGTACACTACTG 720  
 DB 661 GCGGAGGGCTGTGCCAGCCACGGGCTGGACATCACCGGGTGTACACTACTG 720  
 CC and in drug screening for compounds which act as trypsin inhibitors,  
 CC antagonists, agonists, etc. RAS2075\_RNS20790 encode for recombinant  
 CC human beta-II trypsin active site mutants.  
 XX Sequence 771 BP; 152 A; 255 C; 241 G; 123 T; 0 other;

QY 721 GACTGGATCCACCACTATGTCGCCAAAGCCGTGAGGGGCCCGTGT 771  
 DB 721 GACTGGATCCACCACTATGTCGCCAAAGCCGTGAGGGGCCCGTGT 771  
 CC and in drug screening for compounds which act as trypsin inhibitors,  
 CC antagonists, agonists, etc. RAS2075\_RNS20790 encode for recombinant  
 CC human beta-II trypsin active site mutants.  
 XX Sequence 771 BP; 152 A; 255 C; 241 G; 123 T; 0 other;

RESULT 9

AAS20778 AACAAAGAAGATTCGTGGGGTCAGGAGCCAGGGCTGG 660  
 ID AACAAAGAAGATTCGTGGGGTCAGGAGCCAGGGCTGG 660  
 XX AACAAAGAAGATTCGTGGGGTCAGGAGCCAGGGCTGG 660  
 AC AACAAAGAAGATTCGTGGGGTCAGGAGCCAGGGCTGG 660  
 XX AACAAAGAAGATTCGTGGGGTCAGGAGCCAGGGCTGG 660  
 DT 09-APR-2002 (first entry)

DE DNA encoding human beta-II trypsin active site mutant S194A #2.

KW Human; proteolytic trypsinase; protease; recombinant beta-II trypsin;  
 enzyme; mutant; ds.

XX Homo sapiens.

OS Synthetic.

XX

QY 1 GGGCCCTCGAGAARAAGATTCGTGGGGTCAGGAGCCAGGGCTGG 60  
 DB 1 GGGCCCTCGAGAARAAGATTCGTGGGGTCAGGAGCCAGGGCTGG 60  
 XX 1 GGGCCCTCGAGAARAAGATTCGTGGGGTCAGGAGCCAGGGCTGG 60  
 AC 1 GGGCCCTCGAGAARAAGATTCGTGGGGTCAGGAGCCAGGGCTGG 60  
 XX 1 GGGCCCTCGAGAARAAGATTCGTGGGGTCAGGAGCCAGGGCTGG 60  
 DT 120 CAGGTGAGCTGAGAGTCACGGCCATRACTGGATGACTCTCTGGGGCTCCCTCATC 120  
 DE 61 CAGGTGAGCTGAGAGTCACGGCCATRACTGGATGACTCTCTGGGGCTCCCTCATC 120  
 XX 121 CACCCCACTGGGTGCTGACCGCAGCTGCGTGGGACCGTCAAGGATCTGC 180  
 OS 121 CACCCCACTGGGTGCTGACCGCAGCTGCGTGGGACCGTCAAGGATCTGC 180  
 XX 121 CACCCCACTGGGTGCTGACCGCAGCTGCGTGGGACCGTCAAGGATCTGC 180  
 DT 180 CACCCCACTGGGTGCTGACCGCAGCTGCGTGGGACCGTCAAGGATCTGC 180  
 DE 181 GCCCTCAGGGTGCACTGGGAGCAGCTTACTACCAAGACCAACTGGGGCTGC 240  
 XX 181 GCCCTCAGGGTGCACTGGGAGCAGCTTACTACCAAGACCAACTGGGGCTGC 240  
 AC 181 GCCCTCAGGGTGCACTGGGAGCAGCTTACTACCAAGACCAACTGGGGCTGC 240  
 XX 181 GCCCTCAGGGTGCACTGGGAGCAGCTTACTACCAAGACCAACTGGGGCTGC 240  
 DT 241 AGCAGGATCATCGTCACCCACAGTCTCACCCCGGAGACATCGAGCGACATCGCCCTG 300  
 DE 241 AGCAGGATCATCGTCACCCACAGTCTCACCCCGGAGACATCGAGCGACATCGCCCTG 300  
 XX 300 CTGGAGCTGGAGGACCGGGTCAAGGTCTCCAGTCACCGCCAGATCGAGCGACATCGCCCTG 300  
 OS 300 CTGGAGCTGGAGGACCGGGTCAAGGTCTCCAGTCACCGCCAGATCGAGCGACATCGCCCTG 300  
 XX 360 CTGGAGCTGGAGGACCGGGTCAAGGTCTCCAGTCACCGCCAGATCGAGCGACATCGCCCTG 360  
 DT 360 CTGGAGCTGGAGGACCGGGTCAAGGTCTCCAGTCACCGCCAGATCGAGCGACATCGCCCTG 360  
 DE 360 CTGGAGCTGGAGGACCGGGTCAAGGTCTCCAGTCACCGCCAGATCGAGCGACATCGCCCTG 360  
 XX 361 GCCTCAGAGACCTTCCCCGGGGATGGCTGCTGGGTGACTGGCTGGGGATGTGGAC 420  
 OS 361 GCCTCAGAGACCTTCCCCGGGGATGGCTGCTGGGTGACTGGCTGGGGATGTGGAC 420  
 XX 361 GCCTCAGAGACCTTCCCCGGGGATGGCTGCTGGGTGACTGGCTGGGGATGTGGAC 420  
 DT 420 ATATGATGAGGGCTCCACCGCAATTCTCTGAGCAGGGTGAAGGTGCCCATAATGGAA 480  
 DE 420 ATATGATGAGGGCTCCACCGCAATTCTCTGAGCAGGGTGAAGGTGCCCATAATGGAA 480  
 XX 480 ATATGATGAGGGCTCCACCGCAATTCTCTGAGCAGGGTGAAGGTGCCCATAATGGAA 480  
 PI Maffit M, Niles AL, Haak-Frendscho M,  
 XX WPI; 2002-114578/15.  
 DR P-PSDB; AU12012.

PT DNA construct for producing enzymatically-inactive proteolytic trypsinase, comprises DNA sequence encoding proteolytic trypsinase having an active site mutation -

XX Claim 7; Page 91-92; 126pp; English.

PT The present invention relates to recombinant human proteolytic trypsinases active site mutants of these trypsinases and the methods for producing these. The method involves the production of a DNA expression construct comprising a promoter operably linked to a secretion signal sequence which is operably linked to a DNA sequence encoding a proteolytic trypsinase with an active site mutation (the construct drives expression of a mature proteolytic trypsinase that lacks enzymatic activity due to the active site mutation, in hosts transformed to contain the construct comprising a promoter operably linked to a secretion signal sequence which is operably linked to a DNA sequence encoding a proteolytic trypsinase. The active site mutants of proteolytic trypsinase provide a tool to investigate the structural and functional properties of the protease and its enzymatic activity, and for modelling studies. The enzymatically active, recombinant proteolytic trypsinase produced are useful as an antigen to generate anti-human trypsin antibodies

CC and in drug screening for compounds which act as trypsin inhibitors,  
 CC antagonists, agonists, etc. RAS2075\_RNS20790 encode for recombinant  
 CC human beta-II trypsin active site mutants.  
 XX Sequence 771 BP; 152 A; 255 C; 241 G; 123 T; 0 other;

QY 1 GGGCCCTCGAGAARAAGATTCGTGGGGTCAGGAGCCAGGGCTGG 60  
 DB 1 GGGCCCTCGAGAARAAGATTCGTGGGGTCAGGAGCCAGGGCTGG 60  
 CC and in drug screening for compounds which act as trypsin inhibitors,  
 CC antagonists, agonists, etc. RAS2075\_RNS20790 encode for recombinant  
 CC human beta-II trypsin active site mutants.  
 XX Sequence 771 BP; 152 A; 255 C; 241 G; 123 T; 0 other;

QY 61 CAGGTGAGCTGAGAGTCACGGCCATRACTGGATGACTCTCTGGGGCTCCCTCATC 120  
 DB 61 CAGGTGAGCTGAGAGTCACGGCCATRACTGGATGACTCTCTGGGGCTCCCTCATC 120  
 CC and in drug screening for compounds which act as trypsin inhibitors,  
 CC antagonists, agonists, etc. RAS2075\_RNS20790 encode for recombinant  
 CC human beta-II trypsin active site mutants.  
 XX Sequence 771 BP; 152 A; 255 C; 241 G; 123 T; 0 other;

QY 121 CACCCCACTGGGTGCTGACCGCAGCTGCGTGGGACCGTCAAGGATCTGC 180  
 DB 121 CACCCCACTGGGTGCTGACCGCAGCTGCGTGGGACCGTCAAGGATCTGC 180  
 CC and in drug screening for compounds which act as trypsin inhibitors,  
 CC antagonists, agonists, etc. RAS2075\_RNS20790 encode for recombinant  
 CC human beta-II trypsin active site mutants.  
 XX Sequence 771 BP; 152 A; 255 C; 241 G; 123 T; 0 other;

QY 181 GCCCTCAGGGTGCACTGGGAGCAGCTTACTACCAAGACCAACTGGGGCTGC 240  
 DB 181 GCCCTCAGGGTGCACTGGGAGCAGCTTACTACCAAGACCAACTGGGGCTGC 240  
 XX 181 GCCCTCAGGGTGCACTGGGAGCAGCTTACTACCAAGACCAACTGGGGCTGC 240  
 AC 181 GCCCTCAGGGTGCACTGGGAGCAGCTTACTACCAAGACCAACTGGGGCTGC 240  
 XX 181 GCCCTCAGGGTGCACTGGGAGCAGCTTACTACCAAGACCAACTGGGGCTGC 240  
 DT 241 AGCAGGATCATCGTCACCCACAGTCTCACCCCGGAGACATCGAGCGACATCGCCCTG 300  
 DE 241 AGCAGGATCATCGTCACCCACAGTCTCACCCCGGAGACATCGAGCGACATCGCCCTG 300  
 XX 300 CTGGAGCTGGAGGACCGGGTCAAGGTCTCCAGTCACCGCCAGATCGAGCGACATCGCCCTG 300  
 OS 300 CTGGAGCTGGAGGACCGGGTCAAGGTCTCCAGTCACCGCCAGATCGAGCGACATCGCCCTG 300  
 XX 360 CTGGAGCTGGAGGACCGGGTCAAGGTCTCCAGTCACCGCCAGATCGAGCGACATCGCCCTG 360  
 DT 360 CTGGAGCTGGAGGACCGGGTCAAGGTCTCCAGTCACCGCCAGATCGAGCGACATCGCCCTG 360  
 DE 360 CTGGAGCTGGAGGACCGGGTCAAGGTCTCCAGTCACCGCCAGATCGAGCGACATCGCCCTG 360  
 XX 361 GCCTCAGAGACCTTCCCCGGGGATGGCTGCTGGGTGACTGGCTGGGGATGTGGAC 420  
 OS 361 GCCTCAGAGACCTTCCCCGGGGATGGCTGCTGGGTGACTGGCTGGGGATGTGGAC 420  
 XX 361 GCCTCAGAGACCTTCCCCGGGGATGGCTGCTGGGTGACTGGCTGGGGATGTGGAC 420  
 DT 420 ATATGATGAGGGCTCCACCGCAATTCTCTGAGCAGGGTGAAGGTGCCCATAATGGAA 480  
 DE 420 ATATGATGAGGGCTCCACCGCAATTCTCTGAGCAGGGTGAAGGTGCCCATAATGGAA 480  
 XX 480 ATATGATGAGGGCTCCACCGCAATTCTCTGAGCAGGGTGAAGGTGCCCATAATGGAA 480  
 PI Maffit M, Niles AL, Haak-Frendscho M,  
 XX WPI; 2002-114578/15.  
 DR P-PSDB; AU12012.

PT DNA construct for producing enzymatically-inactive proteolytic trypsinase, comprises DNA sequence encoding proteolytic trypsinase having an active site mutation -

XX Claim 7; Page 91-92; 126pp; English.

PT The present invention relates to recombinant human proteolytic trypsinases active site mutants of these trypsinases and the methods for producing these. The method involves the production of a DNA expression construct comprising a promoter operably linked to a secretion signal sequence which is operably linked to a DNA sequence encoding a proteolytic trypsinase with an active site mutation (the construct drives expression of a mature proteolytic trypsinase that lacks enzymatic activity due to the active site mutation, in hosts transformed to contain the construct comprising a promoter operably linked to a secretion signal sequence which is operably linked to a DNA sequence encoding a proteolytic trypsinase. The active site mutants of proteolytic trypsinase provide a tool to investigate the structural and functional properties of the protease and its enzymatic activity, and for modelling studies. The enzymatically active, recombinant proteolytic trypsinase produced are useful as an antigen to generate anti-human trypsin antibodies

KW Human; proteolytic trypsinase; protease; recombinant beta-II trypsinase;  
 KW enzyme; mutant; ds.  
 XX 361 GCCCTCAGAGACTTCGCCCGGGATGCCGCTGGGTCACTGGCTGGCCATGGAC 420  
 OS Homo sapiens.  
 OS Synthetic.  
 XX WO200198470-A2.  
 XX 361 GCCTCAGAGACTTCGCCCGGGATGCCGCTGGGTCACTGGCTGGCCATGGAC 420  
 PD 27-DEC-2001.  
 XX 421 ATATGATGACGCCCTCCACCGCCATTCTCTGAAAGCAGGTGAGCTGGGATGGAC 480  
 PF 20-JUN-2001; 2001WO-US19681.  
 XX 421 ATATGATGACGCCCTCCACCGCCATTCTCTGAAAGCAGGTGAGCTGGGATGGAC 480  
 PR 21-JUN-2000; 2000US-0598982.  
 XX 421 ATATGATGACGCCCTCCACCGCCATTCTCTGAAAGCAGGTGAGCTGGGATGGAC 480  
 PA (PROM-) PROMEGA CORP.  
 XX 481 ACCACATTTGAGCAATACACCCATTGGCCCTACACGGAGACGACGCGCAGTC 540  
 PI Maffit M, Niles AL, Haak-Frendscho M;  
 XX 481 ACCACATTTGAGCAATACACCCATTGGCCCTACACGGAGACGACGCGCAGTC 540  
 DR WPI; 2002-114578/15.  
 DR P-PSDB; AAU12019.  
 XX 541 GTCGGTGAACGACATGCTGTCGGGGACACCCGGAGAACACGGAGGAGACGACGCGCAGTC 540  
 PT DNA construct for producing enzymatically-inactive proteolytic trypsinase, comprises DNA sequence encoding proteolytic trypsinase having an active site mutation -  
 PT 541 GTCGGTGAACGACATGCTGTCGGGGACACCCGGAGGAGACGACGCGCAGTC 540  
 PS Claim 7; Page 110-111; 126pp; English.  
 CC The present invention relates to recombinant human proteolytic trypsinase, active site mutants of these trypsinases and the methods for producing these. The method involves the production of a DNA expression construct comprising a promoter operably linked to a secretion signal sequence which is operably linked to a DNA sequence encoding a proteolytic trypsinase with an active site mutation (the construct drives expression of a mature proteolytic trypsinase that lacks enzymatic activity due to the active site mutation, in hosts transformed to contain the construct). The method is useful for producing enzymatically-active beta-II trypsinase. The active site mutants of proteolytic trypsinase provide a tool to investigate the structural and functional properties of the protease and its enzymatic activity, and for modelling studies. The enzymatically-active, recombinant proteolytic trypsinase produced are useful as an antigen to generate anti-human trypsinase antibodies and in drug screening for compounds which act as trypsin inhibitors, antagonists, agonists, etc. ASN0775-ASN0790 encode for recombinant human beta-II trypsinase active site mutants.  
 CC Sequence 771 BP; 150 A; 256 C; 240 G; 125 T; 0 other;  
 CC Query Match Best local Similarity 98.6%; Score 733.4; DB 24; length 771; Matches 760; Conservative 0; Mismatches 11; Indels 0; Gaps 0; OS Homo sapiens.  
 QY 1 GGGCCCTCGAGAAAGATCGTGCGGGTCAGAGGCCCGAGGAAAGTGCCCGGG 60  
 DB 1 GGGCCCTCGAGAAAGATCGTGCGGGTCAGAGGCCCGAGGAAAGTGCCCGGG 60  
 QY 61 CAGGTGACCTGAGTCACCGGCCCATCTGGATGACTTCGGGGGCTCCCTCATC 120  
 DB 61 CAGGTGACCTGAGTCACCGGCCCATCTGGATGACTTCGGGGGCTCCCTCATC 120  
 QY 121 CACCCCCAAGTGGGCTGTAACCGCCGCGCGTGGGACCGAACGTCAGAGATCTGGCC 180  
 DB 121 CACCCCCAAGTGGGCTGTAACCGCCGCGCGTGGGACCGAACGTCAGAGATCTGGCC 180  
 QY 181 GCCCTCAGGGTGCACCTGGGAGCACCTCTACTACCGACGACCTGGCTGGGACGGACGTCAGAGATCTGGCC 240  
 DB 181 GCCCTCAGGGTGCACCTGGGAGCACCTCTACTACCGACGACCTGGCTGGGACGGACGTCAGAGATCTGGCC 240  
 QY 241 AGCAGGATCATCGTCAGCCACAGTCTACACGCCCGAGTCGGACGACATGCCCTG 300  
 DB 241 AGCAGGATCATCGTCAGCCACAGTCTACACGCCCGAGTCGGACGACATGCCCTG 300  
 QY 301 CTGGAGCTGGAGGCCCGTGAACTGCTCCAGCACCGTCACCGTCACCCCTGGCCCT 360  
 DB 301 CTGGAGCTGGAGGCCCGTGAACTGCTCCAGCACCGTCACCGTCACCCCTGGCCCT 360  
 QY 361 GCCCTCAGAGACTTCGCCCGGGATGCCGCTGGGTCACTGGCTGGCCATGGAC 420  
 PI Maffit M, Niles AL, Haak-Frendscho M;  
 XX 421 ATATGATGACGCCCTCCACCGCCATTCTCTGAAAGCAGGTGAGCTGGGATGGAC 480  
 PR 2002-114578/15.  
 XX 421 ATATGATGACGCCCTCCACCGCCATTCTCTGAAAGCAGGTGAGCTGGGATGGAC 480  
 PA (PROM-) PROMEGA CORP.  
 XX 481 ACCACATTTGAGCAATACACCCATTGGCCCTACACGGAGACGACGCGCAGTC 540  
 DR WPI; 2002-114578/15.  
 DR P-PSDB; AAU12019.  
 XX 541 GTCGGTGAACGACATGCTGTCGGGGACACCCGGAGGAGACGACGCGCAGTC 540  
 PT DNA construct for producing enzymatically-inactive proteolytic trypsinase, comprises DNA sequence encoding proteolytic trypsinase having an active site mutation -  
 PT 541 GTCGGTGAACGACATGCTGTCGGGGACACCCGGAGGAGACGACGCGCAGTC 540  
 PS Claim 7; Page 110-111; 126pp; English.  
 CC The present invention relates to recombinant human proteolytic trypsinases, active site mutants of these trypsinases and the methods for producing these. The method involves the production of a DNA expression construct comprising a promoter operably linked to a secretion signal sequence which is operably linked to a DNA sequence encoding a

proteolytic trypsinase with an active site mutation (the construct drives expression of a mature proteolytic trypsinase that lacks enzymatic activity due to the active site mutation, in hosts transformed to contain the construct). The method is useful for producing enzymatically-active beta-II trypsinase. The active site mutants of proteolytic trypsinase provide a tool to investigate the structural and functional properties of the protease and its enzymatic activity, and for modelling studies. The enzymatically-active, recombinant proteolytic trypsinase produced are useful as an antigen to generate anti-human trypsinase antibodies and in drug screening for compounds which act as trypsinase inhibitors, antagonists, agonists, etc. AAS2075-AAS20790 encode for recombinant human beta-II trypsinase active site mutants.

Sequence 771 BP; 152 A; 256 C; 240 G; 123 T; 0 other;

Query Match 97.7%; Score 753.4; DB 24; Length 771;  
Best Local Similarity 98.6%; Pred. No. 9.6e-145;  
Matches 760; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 1 GGGCCCTCGAGAAAGAATCGTGGGGTCAGGAGGCCAGTGCCCTCG 60  
Db 1 GGGCCCTCGAGAAAGAATCGTGGGGTCAGGAGGCCAGTGCCCTCG 60

Qy 61 CAGGTGAGCTGAGAGTCCACGGCCATACTGGAGTCAGTCTGGGGGCTCCATC 120  
Db 61 CAGGTGAGCTGAGAGTCCACGGCCATACTGGAGTCAGTCTGGGGGCTCCATC 120

Qy 121 CACCCCACTGGGCTGACGCCGCGGCTGGTGGACCGACTCAGGATCTGCC 180  
Db 121 CACCCCACTGGGCTGACGCCGCGGCTGGTGGACCGACTCAGGATCTGCC 180

Qy 181 GCCCTCAGGTGCAACTGGGAGCACCTACTACAGGACAGCTGTCGGTC 240  
Db 181 GCCCTCAGGTGCAACTGGGAGCACCTACTACAGGACAGCTGTCGGTC 240

Qy 241 AGGRAGATCATCGTCACCCACASTCTCACACCCCCAGATCGGAGGCACATGGCCCTG 300  
Db 241 AGGRAGATCATCGTCACCCACAGTCTCACACCCCCAGATCGGAGGCACATGGCCCTG 300

Qy 301 CTGGAGCTGGAGGAGCCGGTGAAGTCTCAGGACAGTGCACCCGCCCT 360  
Db 301 CTGGAGCTGGAGGAGCCGGTGAAGTCTCAGGACAGTGCACCCGCCCT 360

Qy 361 GGCAGAGAACCTCCCGGGATGGCGCTGGACTGCTGGGATGGAC 420  
Db 361 GGCAGAGAACCTCCCGGGATGGCGCTGGACTGCTGGGATGGAC 420

Qy 421 ATATGAGGGCTCCACGGGATTTCTCTGAGCAGGTGAACTCCATATGGAA 480  
Db 421 ATATGAGGGCTCCACGGGATTTCTCTGAGCAGGTGAACTCCATATGGAA 480

Qy 481 AACCCATATTGAGCAGAACATACCCACCTGGCCCTACAGGGAGACAGTGCAGTC 540  
Db 481 AACCCATATTGAGCAGAACATACCCACCTGGCCCTACAGGGAGACAGTGCAGTC 540

Qy 541 GTCCGTGAGCACACTGCTGGCGGGACACCCGGAGGGACTCATGCCAGGGGACTCC 600  
Db 541 GTCCGTGAGCACACTGCTGGCGGGACACCCGGAGGGACTCATGCCAGGGACTCC 600

Qy 601 GGAGGGCTGGTGTGCAAGGATATGGCACCTGGCTCAGGGGGCTGGTCACTGG 660  
Db 601 GGCGGACCAACTGGTGTGCAAGGATATGGCACCTGGCTCAGGGGGCTGGTCACTGG 660

Qy 661 GGCGAGGGCTGGCCCAACGGGCTGGCACCTACCCGGTCACTACTACTGG 720  
Db 661 GGCGAGGGCTGGCCCAACGGGCTGGCACCTACCCGGTCACTACTACTGG 720

Qy 721 GACTGATCCACCACTATGCCAAAGGCCGAGCGGCCGCGCGCT 771  
Db 721 GACTGATCCACCACTATGCCAAAGGCCGAGCGGCCGCGCT 771

AAS2079  
ID AAS2079 standard; DNA; 735 BP.

ID XX

AC XX

ID XX

		PT	an active site mutation -
Db	181 CGGGAGCAGCACCTCTACTACAGAACAGCTGCGCGTCAGAGGATCATGTGCAC	XX	
Qy	259 CCACAGTTCTACACGCCAGATGGAGCCATGCCCTGCTGGAGCTGGAGGCC	PS	Example 1c; Page 115-117; 126pp; English.
Db	241 CCACAGTTCTACACGCCAGATGGAGCCATGCCCTGCTGGAGCTGGAGGCC	XX	
Qy	319 GTGAAGGTCTCACGCCAGTGGAGCCATGCCCTGCTGGAGCTGGAGGCC	CC	The present invention relates to recombinant human proteolytic
Db	301 GTGAGGTCTCACGCCAGTGGAGCCATGCCCTGCTGGAGCTGGAGGCC	CC	tryptases, active site mutants of these tryptases and the methods for
Qy	379 CGGGAGATGCCGCTCTGGTCACTGGCTGGGGATGTGACATGAGCSC	CC	producing these. The method involves the production of a DNA expression
Db	361 CGGGAGATGCCGCTCTGGTCACTGGCTGGGGATGTGACATGAGCSC	CC	construct comprising a promoter operably linked to a secretion signal
Qy	439 CGGCATTTCCTCGAAGCAGGTAAGGCCCCATAATGAAACACATTTG	CC	sequence which is operably linked to a DNA sequence encoding a
Db	421 CGGCATTTCCTCGAAGCAGGTAAGGCCCCATAATGAAACACATTTG	CC	proteolytic tryptase with an active site mutation (the construct drives
Qy	499 AAATACACCTGGGCTACACGGAGACTATGCCAGGAGCTCCGATCG	CC	expression of a mature proteolytic tryptase that lacks enzymatic activity
Db	541 TGTGCCGGAACACCCGGAGGGACTATGCCAGGAGCTCCGATCG	CC	due to the active site mutation, in hosts transformed to contain the
Qy	481 AAATACACCTGGGCTACACGGAGACTATGCCAGGAGCTCCGATCG	CC	construct). The method is useful for producing enzymatically-active
Db	559 TGTGCCGGAACACCCGGAGGGACTATGCCAGGAGCTCCGATCG	CC	beta-II tryptase. The active site mutants of proteolytic tryptase provide
Qy	540 559 TGTGCCGGAACACCCGGAGGGACTATGCCAGGAGCTCCGATCG	CC	a tool to investigate the structural and functional properties of the
Db	541 TGICGCCAACACCCGGAGGGACTATGCCAGGAGCTCCGATCG	CC	protease and its enzymatic activity, and for modelling studies. The
Qy	600 541 TGICGCCAACACCCGGAGGGACTATGCCAGGAGCTCCGATCG	CC	enzymatically-active, recombinant proteolytic tryptase produced are
Db	619 AAATGAGTGTGCACTGGCTACACGGGTGACCTACTTGGA	CC	useful as an antigen to generate anti-human tryptase antibodies
Qy	678 619 AAATGAGTGTGCACTGGCTACACGGGTGACCTACTTGGA	CC	and in drug screening for compounds which act as tryptase inhibitors,
Db	601 AAGCTGAACTGGCTGGGCTGCTAGCTGGGGAGGGCTGCCAG	CC	antagonists, agonists, etc. RAS2075-RAS20790 encode for recombinant
Qy	660 601 AAGCTGAACTGGCTGGGCTGCTAGCTGGGGAGGGCTGCCAG	CC	human beta-II tryptase active site mutants.
Db	679 CCCACCCGGCTGGCATCTACACGGGTGACCTACTTGGA	XX	Sequence 735 BP; 139 A; 246 C; 230 G; 120 T; 0 other;
Qy	738 679 CCCACCCGGCTGGCATCTACACGGGTGACCTACTTGGA	XX	Query Match 95.1%; Score 733.4; DB 24; Length 735;
Db	661 CCCACCCGGCTGGCATCTACACGGGTGACCTACTTGGA	XX	Best Local Similarity 99.9%; Pred. No. 1.1e-140; Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy	720 661 CCCACCCGGCTGGCATCTACACGGGTGACCTACTTGGA	XX	Db 1 ATCGTCGGGGTCAGGAGGCCCTGGAGGAAGTGGCCCTGCGAGGTGAGAGTC
Db	739 GTCCCAAAGGCC 735	XX	Qy 19 ATCGTCGGGGTCAGGAGGCCCTGGAGGAAGTGGCCCTGCGAGGTGAGAGTC
Qy	735 739 GTCCCAAAGGCC 735	XX	Db 79 CACGCCCATACTGATGACTTCGCGGGCCTCCATCCACCCAGTGGTGC
Db	721 GTCCCAAAGGCC 735	XX	Qy 139 ACCGGCGGGGTGCTGGGACCGACTGGATCTGCCCTCAGGGTCACTG
RESULT 13	AAS20787	XX	Db 121 ACCGGCGGGTGGACCGAGATGCGCCCTCAGGGTCACTG
ID	AAS20787 standard; DNA; 735 BP.	XX	Qy 199 CGGGAGCAGCACCTACTACAGAACAGCTGCGCTCAGGGACACTG
AC	AAS20787;	XX	Db 181 CGGGAGCAGCACCTACTACAGAACAGCTGCGCTCAGGGAC
XX	09-APR-2002 (first entry)	XX	Qy 259 CCACAGTTCTCACACGCCAGATGGAGGCCATGCCCTGCGAGGTGGAGGCC
DT	DE DNA encoding human beta-II tryptase active site mutant H44A #4.	XX	Db 241 CCACAGTTCTCACACGCCAGATGGAGGCCATGCCCTGCGAGGTGGAGGCC
XX	KW Human; proteolytic tryptase; protease; recombinant beta-II tryptase; enzyme; mutant; ds.	XX	Qy 319 GTGAAGGTCTCACACGCCAGATGGAGGCCATGCCCTGCGAGGTGGAGGCC
XX	OS Homo sapiens.	OS	Db 301 GTGAAGGTCTCACACGCCAGATGGAGGCCATGCCCTGCGAGGTGGAGGCC
OS	Synthetic.	XX	Qy 379 CGGGAGATGCCGCTGGTCACTGGCTGGGCGATGTGAGCAATGAGC
XX	PN WO200198470-A2.	XX	Db 361 CGGGAGATGCCGCTGGTCACTGGCTGGGCGATGTGAGCAATGAGC
XX	PR 21-JUN-2000; 2000US-0598982.	XX	Qy 439 CGGCATTTCCTCTGAAGCAGGTAAGGCCCCATAATGAAACACATTTG
XX	PD 27-DEC-2001.	XX	Db 421 CGGCATTTCCTCTGAAGCAGGTAAGGCCCCATAATGAAACACATTTG
PA (PROM-) PRIMEGA CORP.	XX	Qy 559 TGTGCCGGAACACCCGGAGGGACTATGCCAGGGACTCCGGAGGCC	
XX	PI Maffit M, Niles AL, Haak-Frendscho M;	XX	Db 541 TGTGCCGGAACACCCGGAGGGACTATGCCAGGGACTCCGGAGGCC
XX	PR WPI; 2002-114578/15.	XX	Qy 619 AGGTGAATGGCACCTGGCTGAGGGGGCTGGGGCTGGGGCTGG
XX	PT P-PSDB; AU12021.	XX	Db 601 AGGTGAATGGCACCTGGCTGAGGGGGCTGGGGCTGGGGCTGG
PT	DNA construct for producing enzymatically-inactive proteolytic tryptase, comprises DNA sequence encoding proteolytic tryptase having	Qy 660 AGGTGAATGGCACCTGGCTGAGGGGGCTGGGGCTGGGGCTGG	
Qy 679 CCCAACGGCCGCTACACCCGGTGTGACCTACTTGGA	XX	Db 678 CCCAACGGCCGCTACACCCGGTGTGACCTACTTGGA	



FH Key location/Qualifiers  
 FT CDS 2..826  
 FT /\*tag= "Human mast cell tryptase II/beta"  
 /product= "the start codon is not indicated"  
 /note= "the start codon is not indicated"  
 PN W09833812-A1.  
 XX XX  
 PD 06-AUG-1998.  
 XX XX  
 PR 30-JAN-1998; 98WO-US01865.  
 XX XX  
 PR 05-FEB-1997; 97US-0037090.  
 PA (BGHM ) BRIGHAM & WOMENS HOSPITAL.  
 XX XX  
 PT Huang, C., Stevens RL;  
 XX DR WPI, 1998-437390/37.  
 DR P-PSDB; AAW63175.  
 PT Tryptase-6 complex inhibitory peptides - used to treat mast cell-mediated inflammatory disorders e.g. asthma  
 XX PS disclosure; Pages 47-48; 69pp; English.  
 CC This cDNA encodes the human mast cell tryptase II/beta which is a homologue of the mouse mast cell protease (mMCP-6) zymogen. The invention provides sequences shown in AAW63160 to AAW63169 that are inhibitors of mMCP-6. These peptides which are tryptase-6 complex inhibitors, can be used for treating a mast cell-mediated inflammatory disorder. The inhibitors can be used to treat inflammatory disorders including asthma, allergic rhinitis, urticaria and anti-oedema, eczematous dermatitis (atopic dermatitis), anaphylaxis, hyperproliferative skin disease, peptic ulcers, inflammatory bowel disorder, hyperresponsiveness and inflammatory skin conditions.  
 CC Sequence 1128 BP; 190 A; 409 C; 329 G; 200 T; 0 other:  
 CC Best Local Similarity 94.8%; Score 731; DB 19; Length 1128;  
 CC Matches 737; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
 QY 17 GAATCGTGGGGTCAGGAGGCCAGAGCAGTGCCCTGGAGGTGAGCTGAGAG 76  
 Db 87 GCATCGTGTGGGGTCAGGAGGCCAGAGCAGTGCCCTGGAGGTGAGCTGAGAG 146  
 QY 77 TCCACGGCCCATACTGGAGCACTTCTGGGGGCTCCCTCACACCCCCAGTGGTGC 136  
 Db 147 TCCACGGCCCATACTGGAGCACTTCTGGGGGCTCCCTCACACCCCCAGTGGTGC 206  
 QY 137 TGACCGCCGGGGTGGCTGGAGCAGGAGCTAGAGAATCTGCCGCCCTCAGGTGAA 196  
 Db 207 TGCACCGCCAGCTGCGGGCTGGAGCAGGAGCTGGATCTGGTGCAC 266  
 QY 197 TGGGGAGGAGCAGCACCTACTACAGGACAGCTGCTGGGGTCAGAGGATCATGGTC 256  
 Db 267 TGGGGAGGAGCACCTACTACAGGACAGCTGCTGGGGTCAGAGGATCATGGTC 326  
 QY 257 ACCCACACTTCTACACCCCGAGATCGAGGGACATCGCCCTGCTGAGCTGGAGC 316  
 Db 327 ACCCACAGTTCTACACCCCGAGATCGAGGGACATCGCCCTGCTGAGCTGGAGC 386  
 QY 317 CGGTGAAGGTCTCCAGGACGTCACCGGTGACCCCTGCTCAGAGACCTTCC 376  
 Db 387 CGGTGAAGGTCTCCAGGACGTCACCGGTGACCCCTGCTCAGAGACCTTCC 446  
 QY 377 CCCCGGGATGGCTGCTGGGTCACTGGCTGGGAGATGGACATGAGCCCTCC 436  
 Db 447 CCCCGGGATGGCTGCTGGGTCACTGGCTGGGAGATGGACATGAGCCCTCC 506  
 QY 437 CACCGCCATTCTCTGAGCAGGTGAAGGTGACCCATATGAAAACACATTGGTACG 496

Search completed: January 31, 2003, 05:24:05  
 Job time : 274 secs

Db 507 CACCGCCATTCTCTGAGCAGGTGAAGGTCCCATTATGAAAACACATTGGTACG 566  
 QY 497 CAAATACACCTTGGGCCCTACAGGGAGAGCAGTCGCAGTCGGTGAGCACATG 556  
 Db 567 CAAATAATACCACCTTGGGCCCTACAGGGAGAGCAGTCGCAGTCGGTGAGCACATG 626  
 QY 557 TGTGTGGGGACACCCGGACTCATGCCAGGGGACTCCGGAGGGCCCTGGT 616  
 Db 627 TGTGTGGGGACACCCGGACTCATGCCAGGGGACTCCGGAGGGCCCTGGT 686  
 QY 617 GCAGAGTGAATGSCACCTGGCTCAGGGGGCTGGTCACTACTTGACTGGATCCACACT 736  
 Db 687 GCAGAGTGAATGSCACCTGGCTCAGGGGGCTGGTCACTACTTGACTGGATCCACACT 806  
 QY 677 AGCCCAAACGGGCCCTGGCATCTACACCCGTGTCACCTACTTGACTGGATCCACACT 736  
 Db 747 AGCCCAAACGGGCCCTGGCATCTACACCCGTGTCACCTACTTGACTGGATCCACACT 833  
 QY 737 ATGTCCCCCAAAAGCCGGTGAAGGGGCC 763  
 Db 807 ATGTCCCCCAAAAGCCGGTGAAGTCAGGC 833







QY 481 AACCACATTTGAGGCGAAATACCCACTGGGCCCTCACGGGAGACGACGCCATC 540  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 481 AACCACATTTGAGGCGAAATACCCACTGGGCCCTCACGGGAGACGACGCCATC 540  
 QY 541 GTCCGTGAGCACATGCTGTCGGGAGAACACCGGAGGGACTCATGCCAGGGACTCC 600  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 541 GTCCGTGAGCACATGCTGTCGGGAGAACACCGGAGGGACTCATGCCAGGGACTCC 600  
 QY 601 GGAGGCCCTGGTGTGCAAGGGAATGSCACCTGGCTCAGGGGGCTGGTCAGCTGG 660  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 601 GGAGGCCCTGGTGTGCAAGGGAATGSCACCTGGCTCAGGGGGCTGGTCAGCTGG 660  
 QY 661 GGCAGGGGCTGGCCAGGCCACGGGCTGGCATCTACACCGGTGACTACTACTTG 720  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 661 GGCAGGGGCTGGCCAGGCCACGGGCTGGCATCTACACCGGTGACTACTACTTG 720  
 QY 721 GACTGGATCACCACATACTGCCCAAAGCCGTGAAAGGGGGCGCGTGT 771  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 721 GACTGGATCACCACATACTGCCCAAAGCCGTGAAAGGGGGCGCGTGT 771  
 RESULT 3  
 LOCUS AR165112  
 DEFINITION Sequence 4 from patent US 6274366.  
 VERSION AR165112  
 KEYWORDS .  
 SOURCE unknown.  
 ORGANISM unclassified.  
 REFERENCE 1 (bases 1 to 771)  
 AUTHORS Maffit,M.A., Niles,A.L. and Haak-Frendscho,M.  
 TITLE Enzymatically-active recombinant human .beta.-tryptase and method  
 JOURNAL Patent: US 6274366-A 4 14-AUG-2001;  
 FEATURES Location/Qualifiers  
 SOURCE 1. .771  
 /organism="unknown"  
 BASE COUNT 149 a 256 c 242 g 124 t  
 ORIGIN  
 Query Match 99 %; Score 764 6; DB 6; Length 771;  
 Best Local Similarity 99.5%; Pred. No. 1.6e-116;  
 Matches 767; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 GGCCCCCTCGAGAAAGAATCGTGGGGTCAGGGCCCCAGGAGCAAGTGGCCCTCG 60  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 1 GGCCCCCTCGAGAAAGAATCGTGGGGTCAGGGCCCCAGGAGCAAGTGGCCCTCG 60  
 QY 61 CAGGTGAGCTGAGAGTCAACGGCCCATACTGGATGCACTCTGGGGCTCCATC 120  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 61 CAGGTGAGCTGAGAGTCAACGGCCCATACTGGATGCACTCTGGGGCTCCATC 120  
 QY 121 CACCCCACTGGGCTGACCGACCTCACTCAGGACAGGTCAAGGATGGCC 180  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 121 CACCCCACTGGGCTGACCGACCTCACTCAGGACAGGTCAAGGATGGCC 180  
 QY 181 GCCCTCAGGGCAACTGGGGAGCAGCACCTCACTCAGGACAGGTCAAGGATGGC 240  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 181 GCCCTCAGGGCAACTGGGGAGCAGCACCTCACTCAGGACAGGTCAAGGATGGC 240  
 QY 241 AGCAGATCATCGTCACCCACACTCTCACACCCCAAGATCGGAGCGACATGCCCTG 300  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 241 AGCAGATCATCGTCACCCACACTCTCACACCCCAAGATCGGAGCGACATGCCCTG 300  
 QY 301 CTGGACTGAGGAGGCCGGTGAAGGTCTCCAGCACTGTCACCGTCAAGGTCAAGG 360  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 301 CTGGACTGAGGAGGCCGGTGAAGGTCTCCAGCACTGTCACCGTCAAGG 360  
 QY 361 GCCCTCAGGACCTCCCGGGGATGCGTGTGCGTGTGACTGCTGGGCGATGTGGAC 420  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 361 GCCCTCAGGACCTCCCGGGGATGCGTGTGCGTGTGACTGCTGGGCGATGTGGAC 420  
 QY 421 ATGATGACGCCCTCCACGCCATTCTCTGAGCAGGTGAGGTCCTTAATGAA 480  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 421 ATGATGACGCCCTCCACGCCATTCTCTGAGCAGGTGAGGTCCTTAATGAA 480  
 QY 481 ACCACATTTGAGCAAATACCCACTGGCTCACCGGAGACGACTCC 540  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 481 ACCACATTTGAGCAAATACCCACTGGCTCACCGGAGACGACTCC 540  
 QY 541 GTCCGTGAGCACATGCTGTCGGGAGAACACCGGAGGGACTCATGCCAGGGACTCC 600  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 541 GTCCGTGAGCACATGCTGTCGGGAGAACACCGGAGGGACTCATGCCAGGGACTCC 600  
 QY 601 GGAGGCCCTGGTGTGCAAGGGAATGSCACCTGGCTCAGGGGGCTGGTCAGCTGG 660  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 601 GGAGGCCCTGGTGTGCAAGGGAATGSCACCTGGCTCAGGGGGCTGGTCAGCTGG 660  
 QY 661 GGCAGGGGCTGGCCAGGCCACGGGCTGGCATCTACACCGGTGACTACTACTTG 720  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 661 GGCAGGGGCTGGCCAGGCCACGGGCTGGCATCTACACCGGTGACTACTACTTG 720  
 QY 721 GACTGGATCACCACATACTGCCCAAAGCCGTGAAAGGGGGCGCGTGT 771  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 721 GACTGGATCACCACATACTGCCCAAAGCCGTGAAAGGGGGCGCGTGT 771  
 RESULT 4  
 LOCUS AX347841  
 DEFINITION Sequence 8 from Patent: WO0198470.  
 VERSION AX347841  
 KEYWORDS SOURCE  
 ORGANISM human.  
 JOURNAL Homo sapiens  
 FEATURES Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 TITLE Recombinant proteolytic tryptases, active site mutants thereof, and  
 methods of making same  
 PATENTWO 0198470 A 8 27-DEC-2001;  
 PROMEGA CORPORATION (US)  
 LOCATION/QUALIFIERS  
 SOURCE 1. .771  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 CDS <7..756  
 /note="unnamed protein product"  
 /codon\_start-1  
 /protein\_id="CAD22418\_1"  
 /db\_xref="GI:18495649"  
 /translation="LEKRVGQEAPESKWPMVSLRVHGPWMMHFGGSILHPOWL  
 TAHCVGPNVKDIALRVLQPLRQHLYQDQLPVSRLVPRYTAIGDIALEH  
 EPKVSSVHTVHLPPASSETEPGMPWVWTFGMWDUDERLPPPEPKQKVPIEMH  
 ICDAKIHHAYTDDVRVLRMLCAGIVTRRQSCOGDSGGPLVCKVNTWLOQAVVSW  
 GEQAQPQRPGTVTRVYLPWLDWVHYYVPKP"  
 BASE COUNT 149 a 256 c 242 g 124 t  
 ORIGIN  
 Query Match 99 %; Score 764 6; DB 6; Length 771;  
 Best Local Similarity 99.5%; Pred. No. 1.6e-116;  
 Matches 767; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 GGCCCCCTCGAGAAAGAATCGTGGGGTCAGGGCCCCAGGAGCACTGCCCTG 60  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 1 GGCCCCCTCGAGAAAGAATCGTGGGGTCAGGGCCCCAGGAGCACTGCCCTG 60  
 QY 61 CAGGTGAGCTGAGGAGGCCGGTGAAGGTCTCCAGCACTGTCACCGTCAAGG 360  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 61 CAGGTGAGCTGAGGAGGCCGGTGAAGGTCTCCAGCACTGTCACCGTCAAGG 360  
 QY 121 CACCCCACTGGGCTGACCGACCTCACTCAGGACAGGTCAAGGATGGCC 180  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 121 CACCCCACTGGGCTGACCGACCTCACTCAGGACAGGTCAAGGATGGCC 180  
 QY 181 GCCCTCAGGGCAACTGGGGAGCAGCACCTCACTCAGGACAGGTCAAGGATGGC 240  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 181 GCCCTCAGGGCAACTGGGGAGCAGCACCTCACTCAGGACAGGTCAAGGATGGC 240  
 QY 241 AGCAGATCATCGTCACCCACACTCTCACACCCCAAGATCGGAGCGACATGCCCTG 300  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 241 AGCAGATCATCGTCACCCACACTCTCACACCCCAAGATCGGAGCGACATGCCCTG 300  
 QY 301 CTGGACTGAGGAGGCCGGTGAAGGTCTCCAGCACTGTCACCGTCAAGG 360  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 301 CTGGACTGAGGAGGCCGGTGAAGGTCTCCAGCACTGTCACCGTCAAGG 360  
 QY 361 GCCCTCAGGACCTCCCGGGGATGCGTGTGCGTGTGACTGCTGGGCGATGTGGAC 420  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 361 GCCCTCAGGACCTCCCGGGGATGCGTGTGCGTGTGACTGCTGGGCGATGTGGAC 420  
 QY 421 CACCCCACTGGGCTGACCCGCCGGGGTGGGACCGACGTCAGTCTCCCTG 180

ICDARYHLGATYGDVDRIVRVDMDICAGNTRDRSDCQGDGGPLVKUNGTLWLOAQVSW  
GEGCQPNRIGGTYVYLDWTHYVPKP"

Db	121 CACCCCGCATGGGTGTGACCCGAGCCACTGGTGGGACCGACCTCACGATCTGCC 180
Qy	181 GCCTCAGGGTCACAATGGGGAGAGCAGCACCCTAATACCGAGGACAGCAGCTGGTC 240
Db	181 GCCTCAGGGTCACAATGGGGAGAGCAGCACCCTAATACCGAGGACAGCAGCTGGTC 240
Qy	241 AGCAGGATCATGTGACCCAGTCTACACGCCAGATGGGGACATGCCCTG 300
Db	241 AGCAGGATCATGTGACCCAGTCTACACGCCAGATGGGGACATGCCCTG 300
Qy	301 CTGGAGATGGAGGAGCCGGTGAAGGTCTCCAGGCCACCTTGCCCCCT 360
Db	301 CTGGAGATGGAGGAGCCGGTGAAGGTCTCCAGGCCACCTTGCCCCCT 360
Qy	361 GCCTCAAGACCTTCCCCCGGGAGACCGCGCTGCTGCTGCTGCTGAC 420
Db	361 GCCTCAAGACCTTCCCCCGGGAGACCGCGCTGCTGCTGCTGAC 420
Qy	421 AATGATGAGCCCTCCACCCGCAATACCCACCTGGGCCTACACGGAGAGCAGCTGGAC 480
Db	421 AATGATGAGCCCTCCACCCGCAATACCCACCTGGGCCTACACGGAGAGCAGCTGGAC 480
Qy	481 AACACATTTGAGACGCAAATACCCACCTGGGCCTACACGGAGAGCAGCTGGAC 540
Db	481 AACACATTTGAGACGCAAATACCCACCTGGGCCTACACGGAGAGCAGCTGGAC 540
Qy	541 GTCTCTGAGACATGCTGCTGCGGACACCCGGAGGACTCATGCCGACTCC 600
Db	541 GTCTCTGAGACATGCTGCTGCGGACACCCGGAGGACTCATGCCGACTCC 600
Qy	601 GGAGGGCCCTGGTGTGCAAGGTGATGGCACCTGGCTGGTCACTGGT 660
Db	601 GGAGGGCCCTGGTGTGCAAGGTGATGGCACCTGGCTGGTCACTGGT 660
Qy	661 GCGGGGGCTGTGCCAGCCCACGGCTGGATACACCCGGTGTGACTACTTG 720
Db	661 GCGGGGGCTGTGCCAGCCCACGGCTGGATACACCCGGTGTGACTACTTG 720
Qy	721 GACTGATCACCCTATGCTCCCAAAGCCGTAAGGGCCCGTGTGACTACTTG 771
Db	721 GACTGATCACCCTATGCTCCCAAAGCCGTAAGGGCCCGTGTGACTACTTG 771
<b>RESULT 5</b>	
AX347838	AX347838 Sequence 5 from Patent WO0198470. DNA linear PAT 01-FEB-2002
LOCUS	Sequence 5 from Patent WO0198470.
DEFINITION	
VERSION	AX347838.1 GI:18495645
KEYWORDS	
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrini; Hominidae; Homo.
AUTHORS	Mafit, M., Niles, A. L. and Haak-Frendscho, M.
TITLE	Recombinant proteolytic tryptases, active site mutants thereof, and methods of making same
JOURNAL	Patent: WO 0198470-A 5 27-DEC-2001;
PROMEGA CORPORATION (US)	
FEATURES	Location/Qualifiers
source	1. -771 /organism="Homo sapiens" <7..756 /db-xref="taxon:9606" /note="unnamed protein product" >/protein_id="CAD22417_1" >/db-xref="GI:18495646" >/translation="LEKRLVGGGAPRSKWPWQVSRLRVHPYRNHFCCGSLTHFQWL TAACPGVPSKDLALRYVOLQEDILVYQDILIPSPSRIVHPYRPTAOQGADTALELE EPVNWSSHVHTVTPPMPMCWVIGWGVDDVNDERLPPFPKLKVVKVPIHENH
<b>RESULT 6</b>	
AX347855	AX347855 Sequence 22 from Patent WO0198470. DNA linear PAT 01-FEB-2002
LOCUS	Sequence 22 from Patent WO0198470.
DEFINITION	
VERSION	AX347855.1 GI:18495662
KEYWORDS	
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrini; Hominidae; Homo.

REFERENCE	1	DB	661	SEQUENCE	GCGGAGGGCTGTGCCCAACCGCCCTGGCATCTACACCGTGTACCAACTTG	720
AUTHORS	Maffit,M., Niles,A.L. and Haak-Frendscho,M.	QY	721	DEFINITION	GACTGATCACCACATGTCGCCAAAGCCGGAACACGGCCGCCRG	771
TITLE	Recombinant proteolytic trypases, active site mutants thereof, and methods of making same	DB	721	ACGTGATCACCACATGTCGCCAAAGCCGGAACACGGCCGCCRG	771	PAT 01-FEB-2002
JOURNAL	Patent: WO 0198410-A 22 27-DEC-2001;					
FEATURES	PROMEGA CORPORATION (US)					
source	Location/Qualifiers					
CDS	1. . 771 /organism="Homo sapiens" /db_xref="taxon:9606" <7. . 756 /note="unnamed protein product" /codon_start=1 /protein_id=CAD22421.1"					
BASE COUNT	149 a 256 c 242 g 124 t					
ORIGIN						
Query Match	98.8%; Score 761.4; DB 6; Length 771; Best Local Similarity 99.2%; Pred. No. 5.3e-116; Matches 765; Conservative 0; Mismatches 6; Indels 0; Gaps 0;					
QY	1 GGGCCCTCGAGAAAAGAATCTGCTGGGGTCAGAGGCCAGTGGCTCG 60					
Db	1 GGGCCCTCGAGAAAAGAATCTGCTGGGGTCAGAGGCCAGTGGCTCG 60					
QY	61 CAGGTGAGCTGAGTCACGGCCATACTGGATCTGGGGCTCCCTCATC 120					
Db	61 CAGGTGAGCTGAGTCACGGCCATACTGGATCTGGGGCTCCCTCATC 120					
QY	121 CACCCCACTGGGCTGACCCGGCTGGTGCCTGGGACGGCTAGGATCTGGC 180					
Db	121 CACCCCACTGGGCTGACCCGGCTGGTGCCTGGGACGGCTAGGATCTGGC 180					
QY	181 GCCCTCAGGGTCACTGGGGACAGCACCTCTACTACAGGACACGGTCAAGGATCTGGC 240					
Db	181 GCCCTCAGGGTCACTGGGGACAGCACCTCTACTACAGGACACGGTCAAGGATCTGGC 240					
QY	241 AGCAGATCATCGCACACAGTCTACACGCCAGATCGGGCAATGCCCTCG 300					
Db	241 AGCAGATCATCGCACACAGTCTACACGCCAGATCGGGCAATGCCCTCG 300					
QY	301 CTGGAGCTGGAGACCCGGTGAAGGCTCTCAGCCACGGTACCCCTGCCCT 360					
Db	301 CTGGAGCTGGAGACCCGGTGAAGGCTCTCAGCCACGGTACCCCTGCCCT 360					
QY	361 GCTCTAGAGACCTTCCCCGGGATGCGGTGCGGGTACTGGTGGCAG 420					
Db	361 GCTCTAGAGACCTTCCCCGGGATGCGGTGCGGGTACTGGTGGCAG 420					
QY	421 ATGTGAGGCCATTCCACGCCATTCTCTGAGCAGGTGAAGGCTCATTTGAA 480					
Db	421 ATGTGAGGCCATTCCACGCCATTCTCTGAGCAGGTGAAGGCTCATTTGAA 480					
QY	481 AACCCATTTGAGCAAAATACCCACCTGGCCCTACACGGGAGACAGTGGCATC 540					
Db	481 AACCCATTTGAGCAAAATACCCACCTGGCCCTACACGGGAGACAGTGGCATC 540					
QY	541 GTCCCTGAGACAGCTGCTGGGGACACCCGGAGGACTCATGGCAGGGACTCC 600					
Db	541 GTCCCTGAGACAGCTGCTGGGGACACCCGGAGGACTCATGGCAGGGACTCC 600					
QY	601 GGAGGCCCTGGCTGCAAGGTGAATGGCCTACACGGGAGACAGTGGCATC 540					
Db	601 GGAGGCCCTGGCTGCAAGGTGAATGGCCTACACGGGAGACAGTGGCATC 540					
QY	601 GGAGGCCCTGGCTGCAAGGTGAATGGCCTACACGGGAGACAGTGGCATC 540					
QY	661 GGCAGGGCTGGCCAGCCAAACGGCCCTGGCATCTACACCGTGTACCAACTTG 720					
BASE COUNT	149 a 257 c 241 g 124 t					
ORIGIN						
FEATURES	PROMEGA CORPORATION (US)					
source	Location/Qualifiers					
CDS	1. . 771 /organism="Homo sapiens" /db_xref="taxon:9606" <7. . 756 /note="unnamed protein product" /codon_start=1 /protein_id=CAD22429.1"					
BASE COUNT	149 a 257 c 241 g 124 t					
ORIGIN						
Query Match	98.5%; Score 759.8; DB 6; Length 771; Best Local Similarity 99.1%; Pred. No. 9.7e-116; Matches 764; Conservative 0; Mismatches 7; Indels 0; Gaps 0;					
QY	1 GGGCCCTCGAGAAAAGAATCTGCTGGGGTCAGAGGCCAGTGGCTCG 60					
Db	1 GGGCCCTCGAGAAAAGAATCTGCTGGGGTCAGAGGCCAGTGGCTCG 60					
QY	61 CAGGTGAGCTGAGTCACGGCCATACTGGATCTGGGGCTCCCTCATC 120					
Db	61 CAGGTGAGCTGAGTCACGGCCATACTGGATCTGGGGCTCCCTCATC 120					
QY	121 CACCCCACTGGGCTGACCCGGTGAAGGCTCTCAGCCACGGTACCCCTGCCCT 360					
Db	121 CACCCCACTGGGCTGACCCGGTGAAGGCTCTCAGCCACGGTACCCCTGCCCT 360					
QY	181 GGCCTCAGGGTCACTGGGGACAGTGGCTGGGACACTCTACAGGACAGTGGCATC 540					
Db	181 GGCCTCAGGGTCACTGGGGACAGTGGCTGGGACACTCTACAGGACAGTGGCATC 540					
QY	301 CTGGAGCTGGGGACAGCTGGCTGGGACACTCTACAGGACAGTGGCATC 540					
Db	301 CTGGAGCTGGGGACAGCTGGCTGGGACACTCTACAGGACAGTGGCATC 540					
QY	361 GGCCTCAGGGTCACTGGGGACAGTGGCTGGGACACTCTACAGGACAGTGGCATC 540					
Db	361 GGCCTCAGGGTCACTGGGGACAGTGGCTGGGACACTCTACAGGACAGTGGCATC 540					

QY 421 ANTGATGAGGCCCTCCACCGGCATTTCTCTGAACGGTGAAGGTCAGCCCCATAATGAA 480  
 Db 421 ANTGATGAGGCCCTCCACCGGCATTTCTCTGAACGGTGAAGGTCAGCCCCATAATGAA 480  
 QY 481 AACCACATTTGTGACGCAAATACCACTTGGCCTACACGGAGAACGACTCCGATC 540  
 Db 481 AACCACATTTGTGACGCAAATACCACTTGGCCTACACGGAGAACGACTCCGATC 540  
 QY 541 GTCGCGTGCAGACATGCTGTGCGGGACACCCGGACTCATSCAGGGCAGCTCC 600  
 Db 541 GTCGCGTGCAGACATGCTGTGCGGGACACCCGGACTCATSCAGGGCAGCTCC 600  
 QY 601 GGAGGGCCCTGTCAGGTAATGGCTGCGAGGAGCTGTCACCGGTGTCACCTACTTG 720  
 Db 601 GGAGGGCCCTGTCAGGTAATGGCTGCGAGGAGCTGTCACCGGTGTCACCTACTTG 720  
 QY 661 GGGAGGGCTGTCAGGTAATGGCTGCGAGGAGCTGTCACCGGTGTCACCTACTTG 720  
 Db 661 GGGAGGGCTGTCAGGTAATGGCTGCGAGGAGCTGTCACCGGTGTCACCTACTTG 720  
 QY 721 GACTGGATCCACCACTATGTCGCCAAMAAGCCGTGAAGCGGCCGCGTGT 771  
 Db 721 GACTGGATCCACCACTATGTCGCCAAMAAGCCGTGAAGCGGCCGCGTGT 771

RESULT 8

AX47857 AX347857 Sequence 24 from Patent WO0198470. DNA linear PAT 01-FEB-2002

LOCUS AX347857 Version 1.0. GI:18495664

DEFINITION Human.

REFERENCE Maffit, M., Niles, A.L. and Haak-Frendscho, M.

AUTHORS Recombinant proteolytic trypases, active site mutants thereof, and methods of making same

TITLE Patent: WO 0198470-A 24 27-DEC-2001;

JOURNAL PROMEGA CORPORATION (US)

FEATURES source

1. .771 /organism="Homo sapiens"  
 /db\_xref="Gr:18495665"  
 /note="unnamed protein product"  
 /codon\_start=1

2. .756 /protein\_id="CAD2422\_1"  
 /db\_xref="Gr:18495665"  
 /transl\_table="LEKRIVGQEPDKIADALRVLVOLRSHQHLYDODLIPVSRTIVHPOFYTAGIDALILETEPVKVSSVHTVTPASETFPGMPCWVIGNDVDERLPPPLIKQVKVPIHENHICDAKYHLGAYIGDDVRDVDMCAUNTRIDSCQGDAGGPLVCKVNGTWLQAGVVWSGSGCCAOENRPGTVTRVYLDWHLQPKK"

BASE COUNT 150 a 255 c 241 g 125 t

ORIGIN

Query Match 97.9%; Score 755; DB 6; Length 771;  
 best local similarity 98.7%; PRED. NO. 6e-115;  
 matches 761; conservative 0; mismatches 10; indels 0; gaps 0; origin

QY 1 GGGCCCTCGAGAAAAGAATCGTGGGGTCAGAGGCCAGAGCAAGTGCCCTGG 60  
 Db 1 GGGCCCTCGAGAAAAGAATCGTGGGGTCAGAGGCCAGAGCAAGTGCCCTGG 60  
 QY 61 CAGGTGAGCTGAGACTCACGCCATACTGATGCACTCTGCGGGCTCCTCATC 120  
 Db 61 CAGGTGAGCTGAGACTCACGCCATACTGATGCACTCTGCGGGCTCCTCATC 120  
 QY 121 CACCCCCAGGGCTGTACCGCCGGGGTGCTGGGGACGGAGTAAAGGATCTGGCC 180

Db 121 CACCCCCAGGGCTGTACCGCCGGGGACGGAGTAAAGGATCTGGCC 180  
 QY 181 GCCTCAGGTGCACTCGGGAGCACCTACTACAGGACAGTCGCGGGTC 240  
 Db 181 GCCTCAGGTGCACTCGGGAGCACCTACTACAGGACAGTCGCGGGTC 240  
 QY 241 ACCAGGATCATGTCACCGACAGTCACAGGACATCGGGACGATCGCGCTG 300  
 Db 241 ACCAGGATCATGTCACCGACAGTCACAGGACATCGGGACGATCGCGCTG 300  
 QY 301 CTGGAGCTGGAGCGGTGAGGTCACCCGTCACCTACTTG 360  
 Db 301 CTGGAGCTGGAGCGGTGAGGTCACCCGTCACCTACTTG 360  
 QY 361 GCCTCAGAGACCTCCCGGGATGCGCAGCTGGGGATGTCAG 420  
 Db 361 GCCTCAGAGACCTCCCGGGATGCGCAGCTGGGGATGTCAG 420  
 QY 421 ANTGATGAGGCCCTCCACCGCATTTCCTCTGCAAGGAGTCACCTACTTG 480  
 Db 421 ANTGATGAGGCCCTCCACCGCATTTCCTCTGCAAGGAGTCACCTACTTG 480  
 QY 481 AACCACATTTGTGACGCAAATGGCTGCGAGGAGCTGTCACCGGTGTCACCTACTTG 540  
 Db 481 AACCACATTTGTGACGCAAATGGCTGCGAGGAGCTGTCACCGGTGTCACCTACTTG 540  
 QY 541 GTCGCGTGCAGACATGCTGTGCGGGACACCCGGACTCATSCAGGGCAGCTCC 600  
 Db 541 GTCGCGTGCAGACATGCTGTGCGGGACACCCGGACTCATSCAGGGCAGCTCC 600  
 QY 601 GGAGGGCCCTGTCAGGTAATGGCTGCGAGGAGCTGTCACCGGTGTCACCTACTTG 720  
 Db 601 GGAGGGCCCTGTCAGGTAATGGCTGCGAGGAGCTGTCACCGGTGTCACCTACTTG 720  
 QY 661 GGGAGGGCTGTCAGGTAATGGCTGCGAGGAGCTGTCACCGGTGTCACCTACTTG 720  
 Db 661 GGGAGGGCTGTCAGGTAATGGCTGCGAGGAGCTGTCACCGGTGTCACCTACTTG 720  
 QY 721 GACTGGATCCACCACTATGTCGCCAAMAAGCCGTGAAGCGGCCGCGTGT 771  
 Db 721 GACTGGATCCACCACTATGTCGCCAAMAAGCCGTGAAGCGGCCGCGTGT 771

RESULT 9

AX347859 AX347859 Sequence 26 from Patent WO0198470. DNA linear PAT 01-FEB-2002

LOCUS AX347859 Version 1.0. GI:18495664

DEFINITION Human.

REFERENCE Maffit, M., Niles, A.L. and Haak-Frendscho, M.

AUTHORS Recombinant proteolytic trypases, active site mutants thereof, and methods of making same

TITLE Patent: WO 0198470-A 26 27-DEC-2001;

JOURNAL PROMEGA CORPORATION (US)

FEATURES source

1. .771 /organism="Homo sapiens"  
 /db\_xref="Gr:18495665"  
 /note="unnamed protein product"  
 /codon\_start=1

2. .756 /protein\_id="CAD2423\_1"  
 /db\_xref="Gr:18495667"  
 /transl\_table="LEKRIVGQEPDKIADALRVLVOLRSHQHLYDODLIPVSRTIVHPOFYTAGIDALILETEPVKVSSVHTVTPASETFPGMPCWVIGNDVDERLPPPLIKQVKVPIHENHICDAKYHLGAYIGDDVRDVDMCAUNTRIDSCQGDAGGPLVCKVNGTWLQAGVVWSGSGCCAOENRPGTVTRVYLDWHLQPKK"

BASE COUNT ORIGIN	152 a 255 c 241 g 123 t	GGCCQAOPNRPGIVTRTYLDWIIHVVPKP#	AUTHORS Maffit, M., Niles, A.L. and Haak-Freindscho, M.
Query Match Best Local Similarity Matches	97.9%; Score 755; DB 6; Length 771; 98.7%; Pred. No. 6e-115; 761; Conservative 0; Mismatches 10; Indels 0; Gaps 0;		TITLE Recombinant proteolytic tryptases, active site mutants thereof, and methods of making same
JOURNAL		PATENT: WO 018470-A 40 27-DEC-2001;	PROMEGA CORPORATION (US)
FEATURES source			LOCATION/QUALIFIERS
QY	1 GGGCCCTCTGAGAAAAGAAATCGTGGGGTCAGAGGCCAGGGAAAGTGCCCCCG 60	1. .771 /organism="Homo sapiens" /db_xref="taxon:9606"	
Db	1 GGGCCCTCTGAGAAAAGAAATCGTGGGGTCAGAGGCCAGGGAAAGTGCCCCCG 60	1. .756 /note="unnamed protein product" /codon_start=1 /protein_id="CAD2243_0_1" /db_xref="GI:18495681"	
QY	61 CACCCCACTGGGTCTGACCCGGGGAGGGACTGGTGGGGACGGAGTCAGGAGTCAGG 180	/translation="GPLEKRTVGOAEPRSKNPWOVSRLVHPYWMHFGGSSLIHPW VLTACVCPDKDALLRVLQRQEQLYDOLLPVSRIIHFQFTAQIJDALIE LREPVNVSHVHTVTPASSETPPGMPCWVQGWGDYNDERLUPPPLKQKVPM NHICDAKYHLGATGDDYRIVDRGDAEGLPPLYCVKNGTMQAGV	
Db	121 CACCCCACTGGGTCTGACCCGGGGAGGGACTGGTGGGGACGGAGTCAGGAGTCAGG 180	1. .756 /translation="GPLEKRTVGOAEPRSKNPWOVSRLVHPYWMHFGGSSLIHPW VLTACVCPDKDALLRVLQRQEQLYDOLLPVSRIIHFQFTAQIJDALIE LREPVNVSHVHTVTPASSETPPGMPCWVQGWGDYNDERLUPPPLKQKVPM NHICDAKYHLGATGDDYRIVDRGDAEGLPPLYCVKNGTMQAGV	
QY	181 GCCCITCAGGGTGCACACTGGGGAGGGACGGACTCTACAGGAGCACGTCAGG 240	1. .756 /translation="GPLEKRTVGOAEPRSKNPWOVSRLVHPYWMHFGGSSLIHPW VLTACVCPDKDALLRVLQRQEQLYDOLLPVSRIIHFQFTAQIJDALIE LREPVNVSHVHTVTPASSETPPGMPCWVQGWGDYNDERLUPPPLKQKVPM NHICDAKYHLGATGDDYRIVDRGDAEGLPPLYCVKNGTMQAGV	
Db	181 GCCCITCAGGGTGCACACTGGGGAGGGACGGACTCTACAGGAGCACGTCAGG 240	1. .756 /translation="GPLEKRTVGOAEPRSKNPWOVSRLVHPYWMHFGGSSLIHPW VLTACVCPDKDALLRVLQRQEQLYDOLLPVSRIIHFQFTAQIJDALIE LREPVNVSHVHTVTPASSETPPGMPCWVQGWGDYNDERLUPPPLKQKVPM NHICDAKYHLGATGDDYRIVDRGDAEGLPPLYCVKNGTMQAGV	
BASE COUNT ORIGIN	150 a 256 c 240 g 125 t	SWEGCAOPNRPGIVTRTYLDWIIHVVPKP#	
Query Match Best Local Similarity Matches	97.7%; Score 753.4; DB 6; Length 771; 98.6%; Pred. No. 1.1e-114; 760; Conservative 0; Mismatches 11; Indels 0; Gaps 0;		
QY	1 GGGCCCTCTGAGAAAAGAAATCGTGGGGTCAGAGGCCAGGGAAAGTGCCCCCG 60	1. .756 /translation="GPLEKRTVGOAEPRSKNPWOVSRLVHPYWMHFGGSSLIHPW VLTACVCPDKDALLRVLQRQEQLYDOLLPVSRIIHFQFTAQIJDALIE LREPVNVSHVHTVTPASSETPPGMPCWVQGWGDYNDERLUPPPLKQKVPM NHICDAKYHLGATGDDYRIVDRGDAEGLPPLYCVKNGTMQAGV	
Db	1 GGGCCCTCTGAGAAAAGAAATCGTGGGGTCAGAGGCCAGGGAAAGTGCCCCCG 60	1. .756 /translation="GPLEKRTVGOAEPRSKNPWOVSRLVHPYWMHFGGSSLIHPW VLTACVCPDKDALLRVLQRQEQLYDOLLPVSRIIHFQFTAQIJDALIE LREPVNVSHVHTVTPASSETPPGMPCWVQGWGDYNDERLUPPPLKQKVPM NHICDAKYHLGATGDDYRIVDRGDAEGLPPLYCVKNGTMQAGV	
QY	301 CTGGAGCTGGAGGGAGGGCTGAAGGCTCAGGCCAACGGTACCCCTCCCCCT 360	1. .756 /translation="GPLEKRTVGOAEPRSKNPWOVSRLVHPYWMHFGGSSLIHPW VLTACVCPDKDALLRVLQRQEQLYDOLLPVSRIIHFQFTAQIJDALIE LREPVNVSHVHTVTPASSETPPGMPCWVQGWGDYNDERLUPPPLKQKVPM NHICDAKYHLGATGDDYRIVDRGDAEGLPPLYCVKNGTMQAGV	
Db	301 CTGGAGCTGGAGGGAGGGCTGAAGGCTCAGGCCAACGGTACCCCTCCCCCT 360	1. .756 /translation="GPLEKRTVGOAEPRSKNPWOVSRLVHPYWMHFGGSSLIHPW VLTACVCPDKDALLRVLQRQEQLYDOLLPVSRIIHFQFTAQIJDALIE LREPVNVSHVHTVTPASSETPPGMPCWVQGWGDYNDERLUPPPLKQKVPM NHICDAKYHLGATGDDYRIVDRGDAEGLPPLYCVKNGTMQAGV	
QY	361 GGCCTCAGAGACCTTCCCCCGGGATGGCGTGGTCACTGGTGGAC 420	1. .756 /translation="GPLEKRTVGOAEPRSKNPWOVSRLVHPYWMHFGGSSLIHPW VLTACVCPDKDALLRVLQRQEQLYDOLLPVSRIIHFQFTAQIJDALIE LREPVNVSHVHTVTPASSETPPGMPCWVQGWGDYNDERLUPPPLKQKVPM NHICDAKYHLGATGDDYRIVDRGDAEGLPPLYCVKNGTMQAGV	
Db	361 GGCCTCAGAGACCTTCCCCCGGGATGGCGTGGTCACTGGTGGAC 420	1. .756 /translation="GPLEKRTVGOAEPRSKNPWOVSRLVHPYWMHFGGSSLIHPW VLTACVCPDKDALLRVLQRQEQLYDOLLPVSRIIHFQFTAQIJDALIE LREPVNVSHVHTVTPASSETPPGMPCWVQGWGDYNDERLUPPPLKQKVPM NHICDAKYHLGATGDDYRIVDRGDAEGLPPLYCVKNGTMQAGV	
QY	421 AACCATTTGTGAGCAAATACACCTGGGCTACAGGGAGAGCTGGCCAT 480	1. .756 /translation="GPLEKRTVGOAEPRSKNPWOVSRLVHPYWMHFGGSSLIHPW VLTACVCPDKDALLRVLQRQEQLYDOLLPVSRIIHFQFTAQIJDALIE LREPVNVSHVHTVTPASSETPPGMPCWVQGWGDYNDERLUPPPLKQKVPM NHICDAKYHLGATGDDYRIVDRGDAEGLPPLYCVKNGTMQAGV	
Db	421 AACCATTTGTGAGCAAATACACCTGGGCTACAGGGAGAGCTGGCCAT 480	1. .756 /translation="GPLEKRTVGOAEPRSKNPWOVSRLVHPYWMHFGGSSLIHPW VLTACVCPDKDALLRVLQRQEQLYDOLLPVSRIIHFQFTAQIJDALIE LREPVNVSHVHTVTPASSETPPGMPCWVQGWGDYNDERLUPPPLKQKVPM NHICDAKYHLGATGDDYRIVDRGDAEGLPPLYCVKNGTMQAGV	
QY	481 AACCATTTGTGAGCAAATACACCTGGGCTACAGGGAGAGCTGGCCAT 540	1. .756 /translation="GPLEKRTVGOAEPRSKNPWOVSRLVHPYWMHFGGSSLIHPW VLTACVCPDKDALLRVLQRQEQLYDOLLPVSRIIHFQFTAQIJDALIE LREPVNVSHVHTVTPASSETPPGMPCWVQGWGDYNDERLUPPPLKQKVPM NHICDAKYHLGATGDDYRIVDRGDAEGLPPLYCVKNGTMQAGV	
Db	481 AACCATTTGTGAGCAAATACACCTGGGCTACAGGGAGAGCTGGCCAT 540	1. .756 /translation="GPLEKRTVGOAEPRSKNPWOVSRLVHPYWMHFGGSSLIHPW VLTACVCPDKDALLRVLQRQEQLYDOLLPVSRIIHFQFTAQIJDALIE LREPVNVSHVHTVTPASSETPPGMPCWVQGWGDYNDERLUPPPLKQKVPM NHICDAKYHLGATGDDYRIVDRGDAEGLPPLYCVKNGTMQAGV	
QY	541 GTCGCGTGGAGCACATGCTGTGGTCACTGGCACCCTGGCAGGGCGCTGG 600	1. .756 /translation="GPLEKRTVGOAEPRSKNPWOVSRLVHPYWMHFGGSSLIHPW VLTACVCPDKDALLRVLQRQEQLYDOLLPVSRIIHFQFTAQIJDALIE LREPVNVSHVHTVTPASSETPPGMPCWVQGWGDYNDERLUPPPLKQKVPM NHICDAKYHLGATGDDYRIVDRGDAEGLPPLYCVKNGTMQAGV	
Db	541 GTCGCGTGGAGCACATGCTGTGGTCACTGGCACCCTGGCAGGGCGCTGG 600	1. .756 /translation="GPLEKRTVGOAEPRSKNPWOVSRLVHPYWMHFGGSSLIHPW VLTACVCPDKDALLRVLQRQEQLYDOLLPVSRIIHFQFTAQIJDALIE LREPVNVSHVHTVTPASSETPPGMPCWVQGWGDYNDERLUPPPLKQKVPM NHICDAKYHLGATGDDYRIVDRGDAEGLPPLYCVKNGTMQAGV	
QY	601 GGCGAGGGCTGTGCCAACGCCAACGGGGCTGGCATCTACCCGGTGTCACT 720	1. .756 /translation="GPLEKRTVGOAEPRSKNPWOVSRLVHPYWMHFGGSSLIHPW VLTACVCPDKDALLRVLQRQEQLYDOLLPVSRIIHFQFTAQIJDALIE LREPVNVSHVHTVTPASSETPPGMPCWVQGWGDYNDERLUPPPLKQKVPM NHICDAKYHLGATGDDYRIVDRGDAEGLPPLYCVKNGTMQAGV	
Db	601 GGCGAGGGCTGTGCCAACGCCAACGGGGCTGGCATCTACCCGGTGTCACT 720	1. .756 /translation="GPLEKRTVGOAEPRSKNPWOVSRLVHPYWMHFGGSSLIHPW VLTACVCPDKDALLRVLQRQEQLYDOLLPVSRIIHFQFTAQIJDALIE LREPVNVSHVHTVTPASSETPPGMPCWVQGWGDYNDERLUPPPLKQKVPM NHICDAKYHLGATGDDYRIVDRGDAEGLPPLYCVKNGTMQAGV	
QY	661 GGCAGGGCTGTGCCAACGCCAACGGGGCTGGCATCTACCCGGTGTCACT 771	1. .756 /translation="GPLEKRTVGOAEPRSKNPWOVSRLVHPYWMHFGGSSLIHPW VLTACVCPDKDALLRVLQRQEQLYDOLLPVSRIIHFQFTAQIJDALIE LREPVNVSHVHTVTPASSETPPGMPCWVQGWGDYNDERLUPPPLKQKVPM NHICDAKYHLGATGDDYRIVDRGDAEGLPPLYCVKNGTMQAGV	
Db	661 GGCAGGGCTGTGCCAACGCCAACGGGGCTGGCATCTACCCGGTGTCACT 771	1. .756 /translation="GPLEKRTVGOAEPRSKNPWOVSRLVHPYWMHFGGSSLIHPW VLTACVCPDKDALLRVLQRQEQLYDOLLPVSRIIHFQFTAQIJDALIE LREPVNVSHVHTVTPASSETPPGMPCWVQGWGDYNDERLUPPPLKQKVPM NHICDAKYHLGATGDDYRIVDRGDAEGLPPLYCVKNGTMQAGV	
QY	721 GACTGGATCACCACATGTCCCCAACAAAGCCGTGGAGGGGCCGCGTGTCACT 720	1. .756 /translation="GPLEKRTVGOAEPRSKNPWOVSRLVHPYWMHFGGSSLIHPW VLTACVCPDKDALLRVLQRQEQLYDOLLPVSRIIHFQFTAQIJDALIE LREPVNVSHVHTVTPASSETPPGMPCWVQGWGDYNDERLUPPPLKQKVPM NHICDAKYHLGATGDDYRIVDRGDAEGLPPLYCVKNGTMQAGV	
Db	721 GACTGGATCACCACATGTCCCCAACAAAGCCGTGGAGGGGCCGCGTGTCACT 720	1. .756 /translation="GPLEKRTVGOAEPRSKNPWOVSRLVHPYWMHFGGSSLIHPW VLTACVCPDKDALLRVLQRQEQLYDOLLPVSRIIHFQFTAQIJDALIE LREPVNVSHVHTVTPASSETPPGMPCWVQGWGDYNDERLUPPPLKQKVPM NHICDAKYHLGATGDDYRIVDRGDAEGLPPLYCVKNGTMQAGV	
RESULT 10			
AX347873	AX347873	771 bp DNA linear PAT 01-FEB-2002	
LOCUS	Sequence 40 from Patent WO0198470.		
DEFINITION			
ACCESSION	AX347873		
VERSION	AX347873.1		
KEYWORDS	human.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrini; Hominoidea; Homo.		
REFERENCE	1		
REFERENCE			

QY	721	GACTGATCCACCATATGCCCAAAAGCGTGAAGCGCCGGCGCTGT	771	Db	421	ATTGATGAGCCCTCCACCGCCATTCTCTGAAGCAGGTGAGTCGCCATATGGAA	480
Db	721	GACTGATCCACCATATGCCCAAAAGCGTGAAGCGCCGGCGCTGT	771	QY	481	AACCACATTGTGACGCATAAATACCACTTGGCCTACACGGGAGCACGTCGCATC	540
DEFINITION	Sequence 42 from Patent WO198470.			Db	541	GTCGGTGACGACATGCTGTGCGGGACACCCGGAGGACTATGCCAAGGAGCC	600
ACCESSION	AX347875			Db	541	GTCGGTGACGACATGCTGTGCGGGACACCCGGAGGACTATGCCAAGGAGCC	600
VERSION	AX347875.1			Db	481	AACCACATTGTGACGCATAAATACCACTTGGCCTACACGGGAGCACGTCGCATC	540
KEYWORDS	human.			QY	541	GTCGGTGACGACATGCTGTGCGGGACACCCGGAGGACTATGCCAAGGAGCC	600
SOURCE	Homo sapiens			Db	541	GTCGGTGACGACATGCTGTGCGGGACACCCGGAGGACTATGCCAAGGAGCC	600
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			Db	601	GGAGGCCCCCTGGTGCAGGTGATGSCACCTGGTCAGGGGGGTGGTCAAGCTGG	660
REFERENCE	1			Db	601	GGAGGCCCCCTGGTGCAGGTGATGSCACCTGGTCAGGGGGGTGGTCAAGCTGG	660
AUTHORS	Maffit,M., Niles,A.L. and Haak-Freindscho,M.			QY	601	GGCGGACACTGGTGTGCAAGGATGSCACCTGGTCAGGGGGGTGGTCAAGCTGG	660
TITLE	Recombinant proteolytic tryptases, active site mutants thereof, and methods of making same			Db	661	GGCGAGGCTGTGCCAACGGCCATGGCCTACACCGTGTCACTACTACTTG	720
JOURNAL	Patent: WO 0198470-A 42 27-DEC-2001; PROMEGA CORPORATION (US)			QY	661	GACTGATCCACCATATGCCCAAAAGCGTGAAGCGCCGGCGCTGT	771
FEATURES	source			Db	721	GACTGATCCACCATATGCCCAAAAGCGTGAAGCGCCGGCGCTGT	771
CDS	1. .771 /organism="Homo sapiens" <!.>.756 /note="unnamed protein product" /protein_id="CAD22431.1" /db_xref="GI:1849563." /translation="LEKTRIVSGEAPRSKWPWQVSLRVHPWMMFCGGSILHPQWVLTAA TAHCWQPDVKDLALRVLQRQLYDQQLPVSRLVPOFYTADQADALLE EPVNVSASHVHTVTPPASETFPGPWVQDNDERLPPFLPKVQKPVIMENH ICDAKYIYGAVYGDVDRVDRMLCAHNTTROSCQDAGGPFLVCKVNGTWLQAGVWS GECCAQPNRPGIYTRVYLYDMLIHVYVKKP"			BASE COUNT	152	a 256 c 240 g 123 t	ORIGIN
ORIGIN				RESULT	11	AX347861	
Query Match	97.7%; Score 753.4; DB 6; Length 771;			LOCUS	AX347861	AX347861	
Best Local Similarity	98.6%; Pred. No. 1.e-114; Mismatches 0; Indels 0; Gaps 0;			DEFINITION	Sequence 28 from Patent WO198470.	735 bp	DNA
Matches	780; Conservative			ACCESSION	AX347861		linear
1	GCCCCCTCGAGAAAAGATCGTGGGGTCAGGAGCAGTCGGCCCTG	60		VERSION	AX347861.1	GI:18495668	PAT 01-FEB-2002
Db	GGGCCCTCGAGAAAAGATCGTGGGGTCAGGAGCAGTCGGCCCTG	60		KEYWORDS	human.		
QY	1	GCCCCCTCGAGAAAAGATCGTGGGGTCAGGAGCAGTCGGCCCTG	60	SOURCE	Homo sapiens		
Db	1	GCCCCCTCGAGAAAAGATCGTGGGGTCAGGAGCAGTCGGCCCTG	60	ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
QY	61	CAGGTGAGCTGAGTCACGGCCCAACTGGTCACTTCGGGGCTCCTC	120	REFERENCE	1		
Db	61	CAGGTGAGCTGAGTCACGGCCCAACTGGTCACTTCGGGGCTCCTC	120	AUTHORS	Maffit,M., Niles,A.L. and Haak-Freindscho,M.		
QY	121	CACCCCACTGGGCTGACCGCCACTGGTCACTTCGGGGCTCCTC	180	TITLE	Recombinant proteolytic tryptases, active site mutants thereof, and methods of making same		
Db	121	CACCCCACTGGGCTGACCGCCACTGGTCACTTCGGGGCTCCTC	180	JOURNAL	Patent: WO 0198470-A 23 27-DEC-2001;		
QY	181	GCCTCAAGGTGCAACTGGGAGCAGCACCTACTACAGGACCGTC	240	FEATURES	source		
Db	181	GCCTCAAGGTGCAACTGGGAGCAGCACCTACTACAGGACCGTC	240	CDS	1. .735 /organism="Homo sapiens" <!.>.735 /note="unnamed protein product" /codon_start=1 /protein_id="CAD22424.1" /db_xref="GI:1849563." /translation="IVGGQEPRSKWPWQVSLRVHPWMMFCGGSILHPQWVLTAA CVGPVKDLALRVLQRQLYDQQLPVSRLVPOFYTADQADALLE VSVHVTVTPPASETFPGPWVQDNDERLPPFLPKVQKPVIMENH KYHGLAYIYGAVYGDVDRVDRMLCAHNTTROSCQDAGGPFLVCKVNGTWLQAGVWS GECCAQPNRPGIYTRVYLYDMLIHVYVKKP"		
Db	181	GCCTCAAGGTGCAACTGGGAGCAGCACCTACTACAGGACCGTC	240	BASE COUNT	139	a 245 c 231 g 120 t	ORIGIN
QY	241	AGCAGGATCATCGTACACCGCCACAGTCTACCGCCACATGCCCTG	300	Query Match	95.3%; Score 735; DB 6; Length 735;		
Db	241	AGCAGGATCATCGTACACCGCCACAGTCTACCGCCACATGCCCTG	300	Best Local Similarity	100.0%; Pred. No. 1.e-111; Mismatches 0; Indels 0; Gaps 0;		
QY	301	CTGGAGCTGGAGGCCGCTGAAAGTCCTACGCCACGTCACAGGT	360	Matches	735; Conservative		
Db	301	CTGGAGCTGGAGGCCGCTGAAAGTCCTACGCCACGTCACAGGT	360	1	ATCGTCGGGGTCAGGAGGCCACAGTGGCTGGCAGGTGAGCTG	60	
QY	361	GCCTCAGAGACCTTCCCCGGGATGCCGTGCTGGTACTGGCTGG	420	QY	19	ATCGTCGGGGTCAGGAGGCCACAGTGGCTGGCAGGTGAGCTG	78
Db	361	GCCTCAGAGACCTTCCCCGGGATGCCGTGCTGGTACTGGCTGG	420	Db	1	ATCGTCGGGGTCAGGAGGCCACAGTGGCTGGCAGGTGAGCTG	60
QY	421	AATGATGAGGCGCTCCACGCCATTCCTGAAAGCAGGTCTCC	480	QY	79	CACGCCATACTGGTCAAGGTCTGGGGCTCTCCATCCAACCCCAGTGG	138
Db	421	AATGATGAGGCGCTCCACGCCATTCCTGAAAGCAGGTCTCC	480	Db	61	CACGCCATACTGGTCAAGGTCTGGGGCTCTCCATCCAACCCCAGTGG	120
QY	361	GCCTCAGAGACCTTCCCCGGGATGCCGTGCTGGTACTGGCTGG	420	QY	139	ACCGCGGGCTGGAGGCCACAGTGGCTGGCCCTCAGGTGCAACT	198
Db	361	GCCTCAGAGACCTTCCCCGGGATGCCGTGCTGGTACTGGCTGG	420	Db	121	ACCGCGGGCTGGAGGCCACAGTGGCTGGCCCTCAGGTGCAACT	180

	BASE COUNT	139 a	246 c	230 g	120 t
QY	199	CGGGCAGCACCTCTACTACCAAGGCCAGCTGCTGCCGTCAAGGAGATCATGTGCAC	258		
Db	181	CGGGCAGCACCTCTACTACCAAGGCCAGCTGCTGCCGTCAAGGAGATCATGTGCAC	240		
QY	259	CCACAGTTCTACACGCCAGARTGGACAGCTGGCCCTSGTGACTGGGGACATCGCCCTSGTGACTGGGGAC	318		
Db	241	CCACAGTTCTACACGCCAGARTGGACAGCTGGACAGCTGGGGACATCGCCCTSGTGACTGGGGAC	300		
QY	319	GTGAGGTCACGCCAGCTCCACACGGTCAACCTGCCCTGCCCCATCGAGACCTTCCC	378		
Db	301	GTGAGGTCACGCCAGCTCCACACGGTCAACCTGCCCTGCCCCATCGAGACCTTCCC	360		
QY	379	CGGGGATGCCGTCTGGGTCACTGTGCTGGGCATGAGACATGAGCGCTCCA	438		
Db	361	CGGGGATGCCGTCTGGGTCACTGTGCTGGGCATGAGACATGAGCGCTCCA	420		
QY	439	CGGCCATTCTCTGAGGAGGTCAAGGTCATAATGAAACACATGACGCCA	498		
Db	421	CGGCCATTCTCTGAGGAGGTCAAGGTCATAATGAAACACATGACGCCA	480		
QY	499	AATACACACTTGGCCATACGGGACAGCAGTCCTGGACATGAGCTG	558		
Db	481	AATACACACTTGGCCATACGGGACAGCAGTCCTGGACATGAGCTG	540		
QY	559	TGTCGGGACACCCGGAGGACTCATCCAGGGACTCTGGGGCCCTGGTGTG	618		
Db	541	TGTCGGGACACCCGGAGGACTCATCCAGGGACTCTGGGGCCCTGGTGTG	600		
QY	619	AAGGTGAATGGCACCTGGCTGCAAGGGGACTCGGGACTGGCTG	678		
Db	601	AAGGTGAATGGCACCTGGCTGCAAGGGGACTGGCTG	660		
QY	679	CCACACGCCGCTGACATCACCGGTGACTACTTGACTTGACTGATCCACCACTATGGACTGAGCATGCT	738		
Db	661	CCACACGCCGCTGACATCACCGGTGACTACTTGACTGAGCATGCT	720		
QY	739	GTCCCCAAAAGCCG	753		
Db	721	GTCCCCAAAAGCCG	735		
RESULT	13				
LOCUS	AX347877	735 bp	DNA	linear	PAT 01-FEB-2002
DEFINITION	Sequence 44 from Patent WO0198470.				
VERSION	AX347877				
KEYWORDS	AX347877.1 GI:18495684				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Mafit,M., Niles A.L. and Haak-Frendscho,M. Recombinant proteolytic tryptases, active site mutants thereof, and methods of making same Patent: WO 0198470-A 44 27-DEC-2001; PROMEGA CORPORATION (US)				
FEATURES	Location/Qualifiers				
source	1. .735 /organism="Homo sapiens" /db_xref="taxon:9606" <1..-.735 /note="unnamed protein product" /codon_start=1 /protein_id="C0D2432_1" /db_xref="GI:18495685" /translation="IVGQEAPRSKWPQVSRLRVHPDWMMHFCGGSLIHPQWLTAA CYGPDPDKLALRKVRLVQHLYDOLIPLUPSRILVPHOFYQIAQIALLEPEVN VSSHHWVHTLPASETPPGMPCWYTGVDNERPPFLKQVKVPIVMENHICDA KHLGATFGDDVRVDMCAGNTRRSCQGDGGPLVCKNGTWIQLQAVVSWGECC AQPNRPGIYTRYVYLDWVIRHYVPKKP"				
RESULT	14				
LOCUS	AR080461	1128 bp	DNA	linear	PAT 31-AUG-2000
DEFINITION	Sequence 15 from patent US 5968782.				
ACCESSION	AR080461				
VERSION	AR080461.1 GI:10007196				
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unclassified				
REFERENCE	1 (bases 1 to 1128)				
AUTHORS	Stevens,R.L.				
TITLE	Mast cell protease that cleaves fibrinogen				

JOURNAL	PATENT	US 5968782-A 15 19-OCT-1999;
FEATURES	Location/Qualifiers	
SOURCE	1. .1128 /organism="unknown"	
BASE COUNT	190 a 409 c 329 g 200 t	
ORIGIN		
Query Match	94.8%; Score 731; DB 6; Length 1128;	
Best Local Similarity	98.7%; Pred. No. 5.1e-111; Mismatches 0; Indels 0; Gaps 0;	
Matches	737; Conservative	
Db	87 GCATCGTGGGGTCAGGAGGCCAGGAGCAAGTGAGTGCAGGTGGAGAG 76	
Qy	17 GAATCGTGGGGTCAGGAGGCCAGGAGCAAGTGAGTGCAGGTGGAGAG 146	
Db	147 TCCACGCCATACTGGATGACTTCAGGGGCTCCTCATCCACCCAGTGGTGC 206	
Qy	137 TGCACGCCAGCAGTGGTGCAGGATCTGGCCCTCAGGGTGAC 196	
Db	207 TGACCGCAGCAGTGGTGCAGGATCTGGCCCTCAGGGTGAC 266	
Qy	197 TGGGGGACAGCACCTCTACTACCAAGGACAGTGTCCCAGTCAGGACATCGTGC 256	
Db	267 TGGGGGAGCACCTCTACTACCAAGGACAGTGTCCCAGTCAGGACATCGTGC 326	
Qy	257 ACCCACAGTCTACCCGCCAGATGGAGCAGCAGTGAGCTGGAGAGC 316	
Db	327 ACCCACAGTCTACCCGCCAGATGGAGCAGCAGTGAGCTGGAGAGC 386	
Qy	317 CGGTGAAGGTTCCAGCCACTGACCGTACCCCTGCTCAGAACCTCC 376	
Db	387 CGGTGAAGGTTCCAGCCACTGACCGTACCCCTGCTCAGAACCTCC 446	
Qy	377 CCCGGGGATGCCCTGCTGGTCACTGGTGGGCGATGTGGACATAGTGGCCCTCC 436	
Db	447 CCCGGGGATGCCCTGCTGGTCACTGGTGGGCGATGTGGACATAGTGGCCCTCC 506	
Qy	437 CACGCCATTCCTCTGTGAGCAGTGAAGTCCCATATGGAAACACATTGTGAGC 496	
Db	507 CACGCCATTCCTCTGTGAGCAGTGAAGTCCCATATGGAAACACATTGTGAGC 566	
Qy	497 CAATATACCCCTGGCCTACACGGAGCAGTCGGCATCTGGTGACATGC 556	
Db	567 CAATAATACCCCTGGCCTACACGGAGCAGTCGGCATCTGGTGACATGC 626	
Qy	557 TGTGTGCCGGAAACACCCGGAGGACTATGCCAGGGGACTCTGGGT 616	
Db	627 TGTGTGCCGGAAACACCCGGAGGACTATGCCAGGGGACTCTGGGT 686	
Qy	617 GCAAGGTGATGGACCCCTGGCTGGGGGGCTGGTAGCTGGGGCAGGSGCTG 676	
Db	687 GCAAGGTGATGGACCCCTGGCTGGGGGGCTGGTAGCTGGGGCAGGSGCTG 746	
Qy	677 AGCCCCACGGCCCTGGCCTACACCGCTGACCTACTCTGGACTGATCCACT 736	
Db	747 AGCCCCACGGCCCTGGCCTACACCGCTGACCTACTCTGGACTGATCCACT 806	
Qy	737 ATGTCCTCCAAAAGCCGTGAGCGCC 763	
Db	807 ATGTCCTCCAAAAGCCGTGAGTCAGGC 833	
RESULT 15		
HUMTRY2A	HUMTRY2A 1128 bp mRNA linear PRI 03-AUG-1993	
LOCUS	Human tryptase-II mRNA, 3' end.	
ACCESSION	M33492	
VERSION	M33492.1 GI:339982	
KEYWORDS	serine protease; tryptase-II.	
SOURCE	Human adult skin, cDNA to mRNA.	
ORGANISM	Homo sapiens	
FEATURES		
SOURCE		
BASE COUNT	190 a 409 c 329 g 200 t	
ORIGIN		
Query Match	94.8%; Score 731; DB 9; Length 1128;	
Best Local Similarity	98.7%; Pred. No. 5.1e-111; Mismatches 0; Indels 0; Gaps 0;	
Matches	737; Conservative	
Db	87 GCATCGTGGGGTCAGGAGGCCAGGAGCAAGTGAGTGCAGGTGGAGAG 76	
Qy	17 GAATCGTGGGGTCAGGAGGCCAGGAGCAAGTGAGTGCAGGTGGAGAG 146	
Db	147 TCCACGCCATACTGGATGACTTCAGGGGCTCCTCATCCACCCAGTGGTGC 206	
Qy	137 TGCACGCCAGCAGTGGTGCAGGATCTGGCCCTCAGGGTGAC 196	
Db	207 TGACCGCAGCAGTGGTGCAGGATCTGGCCCTCAGGGTGAC 266	
Qy	197 TGGGGGACAGCACCTCTACTACCAAGGACAGTGTCCCAGTCAGGACATCGTGC 256	
Db	267 TGGGGGAGCACCTCTACTACCAAGGACAGTGTCCCAGTCAGGACATCGTGC 326	
Qy	257 ACCCACAGTCTACCCGCCAGATGGAGCAGCAGTGAGCTGGAGAGC 316	
Db	327 ACCCACAGTCTACCCGCCAGATGGAGCAGCAGTGAGCTGGAGAGC 386	
Qy	317 CGGTGAAGGTTCCAGCCACTGACCGTACCCCTGCTCAGAACCTCC 376	
Db	387 CGGTGAAGGTTCCAGCCACTGACCGTACCCCTGCTCAGAACCTCC 446	
Qy	377 CCCGGGGATGCCCTGCTGGTCACTGGTGGGCGATGTGGACATAGTGGCCCTCC 436	
Db	447 CCCGGGGATGCCCTGCTGGTCACTGGTGGGCGATGTGGACATAGTGGCCCTCC 506	
Qy	437 CACGCCATTCCTCTGTGAGCAGTGAAGTCCCATATGGAAACACATTGTGAGC 496	
Db	507 CACGCCATTCCTCTGTGAGCAGTGAAGTCCCATATGGAAACACATTGTGAGC 566	
Qy	497 CAAATACCCCTGGCCTACACGGAGCAGTCGGCATCTGGTGAGGACATGC 556	
Db	567 CAATAATACCCCTGGCCTACACGGAGCAGTCGGCATCTGGTGAGGACATGC 626	
Qy	557 TGTGTGCCGGAAACACCCGGAGGACTATGCCAGGGGACTCTGGGT 616	

Fri Jan 31 15:47:10 2003

us-09-598-982-20.rge

Page 11

Db 627 TGTGTGGGGACACCCGGAGGACTCTGCCCGGGCACTCGGAGGGCCCTGGT 686  
Qy 617 GCAAGCTGAATGGCACCTGCCTGCTGGCGCGCGTGTCAGCTGGCGAGGGCTGTGCC 676  
Db 687 GCAAGGTGAATGGCACCTGCCTGCGTGAGCTGGCGAGGGCTGTGCC 746  
Qy 677 AGCCCCAACGGCCCTGGCATCTACACCCGGTCACTACTGGACTGGATCCACCT 736  
Db 747 AGCCAACCGGCCCTGGCATCTACACCCGGTCACTACTGGACTGGATCCACCT 806  
Qy 737 ATGTCCCCRAAAAGCGGTGAAGCGCC 763  
Db 807 ATGTCCCCCAAAAGCGTGTAGTCAGGC 833

Search completed: January 31, 2003, 06:12:33  
Job time : 2817 secs

